

SCORE Search Results Details for Application 10564585 and Search Result 20080623_150908_us-10-564-585- 43.rng.

[Score Home Page](#) [Retrieve Application List](#) [SCORE System Overview](#) [SCORE FAQ Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10564585 and Search Result 20080623_150908_us-10-564-585-43.rng.

[Go Back to previous page](#)

GenCore version 6.2.1
Copyright (c) 1993 - 2008 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2008, 19:33:54 ; Search time 1326 Seconds
(without alignments)
22771.267 Million cell updates/sec

Title: US-10-564-585-43
Perfect score: 2797
Sequence: 1 ttccccagcattcgagaaac.....aaaaaaaaagaaaaaggg 2797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9073515 seqs, 5397694045 residues

Total number of hits satisfying chosen parameters: 18147030

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_200711:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*
14: geneseqn2004c:*
15: geneseqn2004d:*
16: geneseqn2005a:*
17: geneseqn2005b:*
18: geneseqn2005c:*
19: geneseqn2006a:*
20: geneseqn2006b:*
21: geneseqn2006c:*
22: geneseqn2007:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2797	100.0	2797	6	ABL62724	Abl62724 Colon ade
2	2797	100.0	2797	6	ABV94033	Abv94033 Breast ca
3	2797	100.0	2797	8	ABX76311	Abx76311 Lung canc
4	2797	100.0	2797	8	ADB72273	Adb72273 Human SOX
5	2797	100.0	2797	9	ADA02535	Ada02535 Human SOX
6	2797	100.0	2797	10	ADH28813	Adh28813 Human chr
7	2797	100.0	2797	12	ACN38485	Acn38485 Tumour-as
8	2797	100.0	2797	16	ADW47984	Adw47984 Human sex
9	2797	100.0	2797	19	AEK60054	Aek60054 Human SOX
10	2797	100.0	2797	22	AEM95802	Aem95802 Human CML
11	2795.4	99.9	2797	9	ADE95783	Ade95783 Human DNA
12	2793	99.9	3233	8	ADB75556	Adb75556 Prostate
13	2765.8	98.9	4467	5	ABV22264	Abv22264 Human pro
14	2765.8	98.9	4467	5	ABV28101	Abv28101 Human pro
15	2765.8	98.9	4600	6	ABK39749	Abk39749 cDNA enco
16	2765.8	98.9	4600	8	ACA03264	Aca03264 Lung canc
17	2765.8	98.9	4600	8	ACA12078	Aca12078 Human lun
18	2765.8	98.9	4600	9	ADH47316	Adh47316 Human lun
19	2765.8	98.9	4600	12	ADJ21235	Adj21235 Human lun
20	2765.8	98.9	4600	19	AEH23893	Aeh23893 Human lun
21	2765.8	98.9	4912	16	ADX05998	Adx05998 Cyclin-de
22	2765.8	98.9	4912	16	ADZ80563	Adz80563 SRY (sex
23	2765.8	98.9	4912	16	AEA04382	Aea04382 Human cDN
24	2765.8	98.9	4912	16	AEH09647	Aeh09647 SRY (sex
25	2765.8	98.9	4912	19	AEJ48951	Aej48951 Human SRY
26	2765.8	98.9	22801	8	ADB72272	Adb72272 Human SOX
27	2765.8	98.9	22801	9	ADE95782	Ade95782 Human SOX
28	2765.8	98.9	22801	9	ADA02534	Ada02534 Human SOX
29	2765.8	98.9	22801	19	AEK60053	Aek60053 Human SOX
30	2761.2	98.7	4467	4	AAH72650	Aah72650 Human cer
31	2752.4	98.4	5892	9	ADF81488	Adf81488 Leukaemia
32	2752.4	98.4	5892	9	ADF81487	Adf81487 Leukaemia
33	2751.6	98.4	2802	7	AFS92546	Afs92546 Human tra
34	1807	64.6	8801	5	AAS45436	Aas45436 Chemicall
35	1807	64.6	8801	6	ABK28285	Abk28285 DNA trans
36	1807	64.6	8801	6	ABL33740	Abl33740 Human imm
c 37	1721.6	61.6	8801	5	AAS45437	Aas45437 Chemicall
c 38	1721.6	61.6	8801	6	ABK28286	Abk28286 DNA trans
c 39	1721.6	61.6	8801	6	ABL33741	Abl33741 Human imm
40	1425	50.9	1425	8	ADB72274	Adb72274 Human SOX
41	1425	50.9	1425	9	ADE95784	Ade95784 Human SOX
42	1425	50.9	1425	9	ADA02536	Ada02536 Human SOX
43	1425	50.9	1425	12	ACH87117	Ach87117 Human gen
44	1425	50.9	1425	19	AEK60055	Aek60055 Human SOX
c 45	1016	36.3	1584	6	ABQ32052	Abq32052 Oligonucl

ALIGNMENTS

RESULT 1

ABL62724

ID ABL62724 standard; DNA; 2797 BP.

XX

AC ABL62724;

XX

DT 11-JUN-2007 (revised)

DT 15-MAY-2002 (first entry)
 XX
 DE Colon adenocarcinoma related gene sequence SEQ ID NO:1061.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 XX
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 DR PC:NCBI; gi36552.

DR PC_ENCPRO:NCBI; gi36553.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 1061; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
CC
CC Revised record issued on 11-JUN-2007 : Enhanced with precomputed
CC information from BOND.
XX
SQ Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 6; Length 2797;
Best Local Similarity 100.0%; Pred. Loc. 0;
Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Qy	61	ACAGCAAATGTCAGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATGTCAGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAACATAAATCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Db	121	GGAACATAAATCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTCCGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTCCGCCGCCGCGCG	240
Qy	241	CGTCTTCCCGTTCCGGCTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Db	241	CGTCTTCCCGTTCCGGCTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGGTGGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGGTGGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCCTCGAGCTGGGAATCGCCTCTCCCCACGCCCGGCTCCACCGCTCCACGGGCG	480
Db	421	CCGGCCTCGAGCTGGGAATCGCCTCTCCCCACGCCCGGCTCCACCGCTCCACGGGCG	480

Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTCATGGTGTGGTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCCTTCATGGTGTGGTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAAGACAGCG	660
Db	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAAGACAGCG	660
Qy	661	ACAAGATCCCTTTCATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGCTGACTACC	720
Db	661	ACAAGATCCCTTTCATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGCTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781	CGGCCGCGCCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTGCGTGGCAGTGGCGGGG	840
Db	781	CGGCCGCGCCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTGCGTGGCAGTGGCGGGG	840
Qy	841	GCGGCCATGGGGCGGCGGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGGCGGTG	900
Db	841	GCGGCCATGGGGCGGCGGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGGCGGTG	900
Qy	901	CGAGTGGCGGCGGCCCAACTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Db	901	CGAGTGGCGGCGGCCCAACTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Qy	961	CGGGCGGCGGCGGCGGTGGGGTTAGCAAAACCGCAGCCAAAGCTCATCTTGGCAGGCGGCG	1020
Db	961	CGGGCGGCGGCGGCGGTGGGGTTAGCAAAACCGCAGCCAAAGCTCATCTTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCGCTCCTTTCGCGCCGAACAGGCGGGG	1080
Db	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCGCTCCTTTCGCGCCGAACAGGCGGGG	1080
Qy	1081	CCGCGCGCCTGCTGCCCTGGGCGCGCGCGCGCAGCACCACTCGCTGTACAAGGCGCGGA	1140
Db	1081	CCGCGCGCCTGCTGCCCTGGGCGCGCGCGCGCAGCACCACTCGCTGTACAAGGCGCGGA	1140
Qy	1141	CTCCAGCGCCTCGGCCCTCCGCTCCTCGGCAGCCTCGGCCCTCCGAGCGCTCGCGGCC	1200
Db	1141	CTCCAGCGCCTCGGCCCTCCGCTCCTCGGCAGCCTCGGCCCTCCGAGCGCTCGCGGCC	1200
Qy	1201	CGGGCAAGCACCTGGCGGAGAAGAAGGTGAAGCGGCTACCTGTTGGCGGCTGGGCA	1260
Db	1201	CGGGCAAGCACCTGGCGGAGAAGAAGGTGAAGCGGCTACCTGTTGGCGGCTGGGCA	1260
Qy	1261	CGTCGTCGTCGCCCTGGGCGGCGTGGGCGCGGGAGGCCGACCCAGCGACCCCTGGGCC	1320
Db	1261	CGTCGTCGTCGCCCTGGGCGGCGTGGGCGCGGGAGGCCGACCCAGCGACCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGGCGCGGGCTGCTCGCCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Db	1321	TGTACGAGGAGGAGGGCGCGGGCTGCTCGCCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Qy	1381	GCGCGGCTCGTCCCCCGCGCGCGGCGCTCGCCCGCGCACCACCGCGGCTACGCCAGCC	1440
Db	1381	GCGCGGCTCGTCCCCCGCGCGCGGCGCTCGCCCGCGCACCACCGCGGCTACGCCAGCC	1440
Qy	1441	TGCGCGCGCCTCGCCCGCCCCGTCCAGCGCGCCTCGCACGCTCCTCCTCGGCCCTCGT	1500
Db	1441	TGCGCGCGCCTCGCCCGCCCCGTCCAGCGCGCCTCGCACGCTCCTCCTCGGCCCTCGT	1500

Qy	1501	CCCACCTCCTCCTCTTCTCTCTCTCGGGCTCCTCGTCTCCGACGACGAGTTCGAAGACG	1560
Db	1501	CCCACCTCCTCCTCTTCTCTCTCTCGGGCTCCTCGTCTCCGACGACGAGTTCGAAGACG	1560
Qy	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Db	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Qy	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCTCCGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCTCCGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGGCGCGCAGGCAGGGAGAAGGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGGCGCGCAGGCAGGGAGAAGGGC	1800
Qy	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Db	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGCTTTTGACCCGCGCTCCCATCCCCCACT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGCTTTTGACCCGCGCTCCCATCCCCCACT	1980
Qy	1981	TCCGGGCGGGGACCCACTCTGCCAGCCGAGGGAGCGGAGGAGGAAGAGGGTAGAC	2040
Db	1981	TCCGGGCGGGGACCCACTCTGCCAGCCGAGGGAGCGGAGGAGGAAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Qy	2101	TTGAGTTTGCTCCCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101	TTGAGTTTGCTCCCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Qy	2161	TTGCTGCACCCCGACGAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161	TTGCTGCACCCCGACGAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Qy	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTGTTCTTCTGGTCAAGAAAGGAGGGGA	2280
Db	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTGTTCTTCTGGTCAAGAAAGGAGGGGA	2280
Qy	2281	GAACCCAGCGCACCCCTCCCCCTTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAACCCAGCGCACCCCTCCCCCTTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Db	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Qy	2401	GCAGGGCAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Db	2401	GCAGGGCAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Qy	2461	AGCCGAGGAGGAGATGTTAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG	2520

```

Db      2461 AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCCGCGCTAGGAAATG 2520
Qy      2521 ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGCGGAGGAGGACA 2580
      |||||
Db      2521 ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGCGGAGGAGGACA 2580
Qy      2581 CGAACTGGAAGGGGTTTACGGTCAAACCTGAAATGGATTTCACACGTTGGGGAGCTGGCGG 2640
      |||||
Db      2581 CGAACTGGAAGGGGTTTACGGTCAAACCTGAAATGGATTTCACACGTTGGGGAGCTGGCGG 2640
Qy      2641 CGGCGGCTGCTGGGCTCCCGCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
      |||||
Db      2641 CGGCGGCTGCTGGGCTCCCGCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Qy      2701 GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
      |||||
Db      2701 GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
Qy      2761 GCGGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797
      |||||
Db      2761 GCGGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797

```

RESULT 2

ABV94033

```

ID      ABV94033 standard; cDNA; 2797 BP.
XX
AC      ABV94033;
XX
DT      11-JUN-2007 (revised)
DT      08-JAN-2003 (first entry)
XX
DE      Breast carcinoma related nucleotide sequence SEQ ID NO:24.
XX
KW      Human; breast carcinoma; cancer; tumour; cytostatic; anti-tumour; gene;
KW      ss.
XX
OS      Homo sapiens.
XX
PN      WO200246467-A2.
XX
PD      13-JUN-2002.
XX
PF      07-DEC-2001; 2001WO-IB002811.
XX
PR      08-DEC-2000; 2000US-0254090P.
PR      07-DEC-2001; 2001US-00007926.
XX
PA      (IPSO-) IPSOGEN.
XX
PI      Bertucci F, Houlgatte R, Birnbaum D, Nguyen C, Viens P, Fert V;
XX
DR      WPI; 2002-619023/66.
DR      PC:NCBI; gi36552.
DR      PC_ENCPRO:NCBI; gi36553.
XX
PT      Novel polynucleotide library useful in molecular characterization of a
PT      carcinoma, comprising a pool of polynucleotide sequences or its
PT      subsequences which are either underexpressed or overexpressed in tumor
PT      cells.
XX
PS      Claim 1; Page 124-125; 401pp; English.
XX
CC      The present invention describes a polynucleotide library (I) useful in
CC      the molecular characterisation of a carcinoma, comprising a pool of
CC      polynucleotides or its subsequences which are either underexpressed or
CC      overexpressed in tumour cells, and correspond to any of the

```

polynucleotide sequences chosen from the 468 sequences given in ABV94010 to ABV94477. Also described: (I) a polynucleotide array (II) useful for the prognosis or diagnostic of tumour, comprising (I); and (2) detecting (MI) differentially expressed polynucleotide sequences which are correlated with a cancer, involves obtaining a polynucleotide sample from a patient, and reacting the polynucleotide sample obtained with a probe immobilised on a solid support, where the probe comprises any combination of the polynucleotide sequences of (I) or its expression products encoded by polynucleotide sequences of (I), and detecting the reaction product. (I) have cytostatic activities and can be used as anti-tumour agents. (I) is useful in molecular characterisation of a carcinoma. (I) and (II) are useful for the prognosis or diagnostic of tumour, in differentiating a normal cell from a cancer cell, detecting a hormone sensitive tumour cell, differentiating a tumour with lymph nodes from a tumour without lymph nodes, differentiating antracycline-sensitive tumours from antracycline-insensitive tumours, and classifying good and poor prognosis primary breast tumours. (I) is useful for large-scale molecular characterisation of breast cancer that help in prediction, prognosis and cancer treatment, and for detecting differentially expressed genes that correlated with a cancer

Revised record issued on 11-JUN-2007 : Enhanced with precomputed information from BOND.

Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 6; Length 2797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Qy	61	ACAGCAAATCGACGCGGTGAGAGAGCGAGAGAGAGAGGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATCGACGCGGTGAGAGAGCGAGAGAGAGAGGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAATAATACTCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT	180
Db	121	GGAATAATACTCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTTCGGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTTCGGCCGCCGCGCG	240
Qy	241	CGTCTTCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCGG	300
Db	241	CGTCTTCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCGG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGGCGCCCGCCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGGCGCCCGCCGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCCTCGAGCTGGGAATCGCCTCCTCCCCACGCCGGCTCCACCGCCTCCACGGGCG	480
Db	421	CCGGCCTCGAGCTGGGAATCGCCTCCTCCCCACGCCGGCTCCACCGCCTCCACGGGCG	480
Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTATGTTGTTGTCGACAGATCGAGCGGCGCAAGATCATGAGCAGTCGCCCGACA	600

Db	541		ACGCCTTCATGGTGTGGTCGCAGATCGAGCGGCCAAGATCATGGAGCAGTCGCCCGGACA	600
Qy	601		TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAAGACAGCG	660
Db	601		TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAAGACAGCG	660
Qy	661		ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGTGACTACC	720
Db	661		ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGTGACTACC	720
Qy	721		CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721		CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781		CGGCCGCCGCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTGCGAGTGGCGGGG	840
Db	781		CGGCCGCCGCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTGCGAGTGGCGGGG	840
Qy	841		GCGGCCATGGGGGCGCGGCGGCGGGAGCAGCAACGCGGGGAGGAGGCGCGGTG	900
Db	841		GCGGCCATGGGGGCGCGGCGGCGGGAGCAGCAACGCGGGGAGGAGGCGCGGTG	900
Qy	901		CGAGTGGCGGCGGCCAACTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Db	901		CGAGTGGCGGCGGCCAACTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Qy	961		CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTTGGCAGGCGGCG	1020
Db	961		CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTTGGCAGGCGGCG	1020
Qy	1021		GCGGCGGCGGGAAAGCAGCGGCTGCCGCCGCCCTCCTTCGCCGCCGAACAGGCGGGG	1080
Db	1021		GCGGCGGCGGGAAAGCAGCGGCTGCCGCCGCCCTCCTTCGCCGCCGAACAGGCGGGG	1080
Qy	1081		CCGCCGCCCTGCTGCCCTGGGCGCGCGGCCACCACTCGCTGTACAAGGCGCGGA	1140
Db	1081		CCGCCGCCCTGCTGCCCTGGGCGCGCGGCCACCACTCGCTGTACAAGGCGCGGA	1140
Qy	1141		CTCCAGCGCCTCGGCTCCGCTCTCGGCAGCCTCGGCTCCGAGCGCTCGCGGCC	1200
Db	1141		CTCCAGCGCCTCGGCTCCGCTCTCGGCAGCCTCGGCTCCGAGCGCTCGCGGCC	1200
Qy	1201		CGGGCAAGCACTTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTCCGCGGCTGGGCA	1260
Db	1201		CGGGCAAGCACTTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTCCGCGGCTGGGCA	1260
Qy	1261		CGTCGTCGTCGCCCGTGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Db	1261		CGTCGTCGTCGCCCGTGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Qy	1321		TGTACGAGGAGGAGGCGCGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCGCGAGCA	1380
Db	1321		TGTACGAGGAGGAGGCGCGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCGCGAGCA	1380
Qy	1381		GCGCCGCTCGTCCCCGCCGCCGCGCTCGCCGCCGACACCGCGGCTACGCGAGCC	1440
Db	1381		GCGCCGCTCGTCCCCGCCGCCGCGCTCGCCGCCGACACCGCGGCTACGCGAGCC	1440
Qy	1441		TGCGCGCGGCTCGCCGCCCGCTCCAGCGCGCCTCGCACGCTCTCTCGGCTCGT	1500
Db	1441		TGCGCGCGGCTCGCCGCCCGCTCCAGCGCGCCTCGCACGCTCTCTCGGCTCGT	1500
Qy	1501		CCCACTCCTCTCTTCTCTCTCTCGGCTCTCTGCTCTCCGACGACGAGTTCAAGACG	1560
Db	1501		CCCACTCCTCTCTTCTCTCTCTCGGCTCTCTGCTCTCCGACGACGAGTTCAAGACG	1560

Qy	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGCAGCTTCAGTT	1620
Db	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGCAGCTTCAGTT	1620
Qy	1621	CGTCGTCGGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTCGGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCCGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCCGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGCGCGCAGGCAGGAGAGAAGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGCGCGCAGGCAGGAGAGAAGGC	1800
Qy	1801	CGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAGGAAAAAGAAAGAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Db	1861	GAGTTTAAAGAGAAAAGGAAAAAGAAAGAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTTGACCCCGCTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTTGACCCCGCTCCCATCCCCACCT	1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGGAGGAGGAAGAGGTAGAC	2040
Db	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGGAGGAGGAAGAGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAGGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAGGCGAC	2100
Qy	2101	TTGAGTTTGCTCCCTTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101	TTGAGTTTGCTCCCTTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Qy	2161	TTGCTGACCCCGCAGCAAGAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161	TTGCTGACCCCGCAGCAAGAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Qy	2221	CTTCTGCATACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Db	2221	CTTCTGCATACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Qy	2281	GAACCCAGCGCACCCCTCCCCCTTTTTTAAACGGGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAACCCAGCGCACCCCTCCCCCTTTTTTAAACGGGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Db	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Qy	2401	GCAGGCGAATTCCCGTTTGGGGCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTGTC	2460
Db	2401	GCAGGCGAATTCCCGTTTGGGGCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTGTC	2460
Qy	2461	AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCAGTGTGACCGGCGCTAGGAAATG	2520
Db	2461	AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCAGTGTGACCGGCGCTAGGAAATG	2520
Qy	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Db	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGCGGGGGCGGAGGAGACA	2580

```

Qy      2581 CGAACTGGAAGGGGGTTCACGGTCAAAC TGAAATGGATTGACGTTGGGGAGCTGGCGG 2640
Db      2581 CGAACTGGAAGGGGGTTCACGGTCAAAC TGAAATGGATTGACGTTGGGGAGCTGGCGG 2640

Qy      2641 CGGCGGCTGCTGGGCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Db      2641 CGGCGGCTGCTGGGCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700

Qy      2701 GACCCCGGAGGCGTGGAGGAGGAGAGACTGTTTGTATGTGGTACAGGGGCAGTCAGTGGAG 2760
Db      2701 GACCCCGGAGGCGTGGAGGAGGAGAGACTGTTTGTATGTGGTACAGGGGCAGTCAGTGGAG 2760

Qy      2761 GGCAGTGGTTTCGGAAAAAAAAAAGAAAAAGGG 2797
Db      2761 GGCAGTGGTTTCGGAAAAAAAAAAGAAAAAGGG 2797

```

RESULT 3

ABX76311

ID ABX76311 standard; DNA; 2797 BP.

XX

AC ABX76311;

XX

DT 11-JUN-2007 (revised)

DT 02-APR-2003 (first entry)

XX

DE Lung cancer-associated polynucleotide #175.

XX

KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX

OS Unidentified.

XX

PN WO200286443-A2.

XX

PD 31-OCT-2002.

XX

PF 18-APR-2002; 2002WO-US012476.

XX

PR 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX

PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX

PI Aziz N, Murray R;

XX

DR WPI; 2003-093161/08.

DR P-PSDB; ABU56582.

DR PC:NCBI; gi36552.

XX

PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX

PS Claim 22; Page 322-323; 453pp; English.

XX

CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological

CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

CC
 CC Revised record issued on 11-JUN-2007 : Enhanced with precomputed
 CC information from BOND.

XX
 SQ Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 8; Length 2797;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Qy	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGGGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGGGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAACATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Db	121	GGAACATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCCAGTTCGGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCCAGTTCGGCCGCCGCGCG	240
Qy	241	CGTCTTCCCGTTCGGCGTGTGCTTGGCCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCG	300
Db	241	CGTCTTCCCGTTCGGCGTGTGCTTGGCCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCTTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGGCTCCACGCCCTCCACGGGCG	480
Db	421	CCGGCTTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGGCTCCACGCCCTCCACGGGCG	480
Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTCATGGTGTGGTTCGAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCGACA	600
Db	541	ACGCCTTCATGGTGTGGTTCGAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCGACA	600

Qy	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAGACAGCG	660
Db	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAGACAGCG	660
Qy	661	ACAAGATCCCTTTTCATTTCGAGAGCGGAGCGGCTGCGCTCAAGCACATGGCTGACTACC	720
Db	661	ACAAGATCCCTTTTCATTTCGAGAGCGGAGCGGCTGCGCTCAAGCACATGGCTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781	CGGCCGCGGCTCCTCAAGCCGGGGAGAAGGGAGACAAGTTCGGTGGCAGTGGCGGGG	840
Db	781	CGGCCGCGGCTCCTCAAGCCGGGGAGAAGGGAGACAAGTTCGGTGGCAGTGGCGGGG	840
Qy	841	GCGGCCATGGGGCGGCGGCGGCGGGAGCAGCAACGCGGGGGAGGAGGCGGCGTG	900
Db	841	GCGGCCATGGGGCGGCGGCGGCGGGAGCAGCAACGCGGGGGAGGAGGCGGCGTG	900
Qy	901	CGAGTGGCGGCGGCGCAACTCCAAACCGGCGCAGAAAAAGAGTTCGGCTCCAAAGTGG	960
Db	901	CGAGTGGCGGCGGCGCAACTCCAAACCGGCGCAGAAAAAGAGTTCGGCTCCAAAGTGG	960
Qy	961	CGGGCGGCGGCGGCGTGGGGTTAGCAAAACCGCAGCGCAAGCTCATCTTGGCAGGCGGCG	1020
Db	961	CGGGCGGCGGCGGCGTGGGGTTAGCAAAACCGCAGCGCAAGCTCATCTTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAGCAGCGGCTGCCGCCGCCCTCCTTCGCGCGGAACAGGCGGGG	1080
Db	1021	GCGGCGGCGGGAAGCAGCGGCTGCCGCCGCCCTCCTTCGCGCGGAACAGGCGGGG	1080
Qy	1081	CCGCCGCCCTGCTGCCCTGGGCGCGCGCGCACCACCTCGTGTACAAGGCGCGGA	1140
Db	1081	CCGCCGCCCTGCTGCCCTGGGCGCGCGCGCACCACCTCGTGTACAAGGCGCGGA	1140
Qy	1141	CTCCAGCGCCTCGGCCTCCGCCTCCTCGGCAGCTCGGCCTCCGAGCGCTCGCGGCC	1200
Db	1141	CTCCAGCGCCTCGGCCTCCGCCTCCTCGGCAGCTCGGCCTCCGAGCGCTCGCGGCC	1200
Qy	1201	CGGGCAAGCACTGGCGGAGAAGAAGGTGAAGCGCGTCTACCTGTTGGCGGCTGGGCA	1260
Db	1201	CGGGCAAGCACTGGCGGAGAAGAAGGTGAAGCGCGTCTACCTGTTGGCGGCTGGGCA	1260
Qy	1261	CGTCGTCGTCGCCGCTGGGCGGCGTGGGCGGGAGCCGACCCAGCAGCCCTGGGCC	1320
Db	1261	CGTCGTCGTCGCCGCTGGGCGGCGTGGGCGGGAGCCGACCCAGCAGCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCTGAGCGGCGCAGCA	1380
Db	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCTGAGCGGCGCAGCA	1380
Qy	1381	GCGCGGCTCGTCCCCGCGCGCGGCGCTCGCCCGCCGACACCGCGGCTACGCCAGCC	1440
Db	1381	GCGCGGCTCGTCCCCGCGCGCGGCGCTCGCCCGCCGACACCGCGGCTACGCCAGCC	1440
Qy	1441	TGCGCGCGGCTCGCCCGCCCGTCCAGCGCGCTCGCACGCGTCTCTCGGCTCGT	1500
Db	1441	TGCGCGCGGCTCGCCCGCCCGTCCAGCGCGCTCGCACGCGTCTCTCGGCTCGT	1500
Qy	1501	CCCACTCTCTCTTCTCTCTCTCGGGCTCTCTGCTCTCCGACGACGAGTTCGAAGACG	1560
Db	1501	CCCACTCTCTCTTCTCTCTCTCGGGCTCTCTGCTCTCCGACGACGAGTTCGAAGACG	1560
Qy	1561	ACCTGCTCGACTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGCGAGTTTCAGTT	1620
Db	1561	ACCTGCTCGACTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGCGAGTTTCAGTT	1620

Qy	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCCGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCCGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACCTACTGAAGGGCGCGCAGGCAGGAGAAGGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACCTACTGAAGGGCGCGCAGGCAGGAGAAGGGC	1800
Qy	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAGAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Db	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAGAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTTGACCCGCGCTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTTGACCCGCGCTCCCATCCCCACCT	1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Db	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Qy	2101	TTGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101	TTGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Qy	2161	TTGCTGCACCCCAAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161	TTGCTGCACCCCAAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Qy	2221	CTTCCTGCATCACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Db	2221	CTTCCTGCATCACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Qy	2281	GAAACCGAGCGACCCCTCCCCCCTTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAAACCGAGCGACCCCTCCCCCCTTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTGCGCGATGGCAGATGTTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Db	2341	GGTGACGAATTGCGCGATGGCAGATGTTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Qy	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Db	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Qy	2461	AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG	2520
Db	2461	AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG	2520
Qy	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Db	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Qy	2581	CGAACTGGAAGGGGTTACAGTCAAAGTGAATGGATTGACAGTTGGGGAGCTGGCGG	2640

```

Db      2581 CGAACTGGAAGGGGGTTCACGCTCAAACCTGAAATGGATTTCACGTTGGGGAGCTGGCGG 2640
Qy      2641 CGGCGGCTGCTGGGCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
        |||
Db      2641 CGGCGGCTGCTGGGCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Qy      2701 GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
        |||
Db      2701 GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
Qy      2761 GGCAGTGGTTTCGGAAAAAAGAAAAAAGGG 2797
        |||
Db      2761 GGCAGTGGTTTCGGAAAAAAGAAAAAAGGG 2797

```

RESULT 4

ADB72273

ID ADB72273 standard; mRNA; 2797 BP.

XX

AC ADB72273;

XX

DT 11-JUN-2007 (revised)

DT 04-DEC-2003 (first entry)

XX

DE Human SOX4 mRNA.

XX

KW human; ss; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;

KW cancer; neoplasm; adenocarcinoma; sarcoma.

XX

OS Homo sapiens.

XX

PN WO2003008583-A2.

XX

PD 30-JAN-2003.

XX

PF 26-DEC-2001; 2001WO-US051291.

XX

PR 02-MAR-2001; 2001US-00798586.

PR 23-OCT-2001; 2001US-00004113.

PR 08-NOV-2001; 2001US-00052482.

PR 30-NOV-2001; 2001US-00997722.

PR 20-DEC-2001; 2001US-00034650.

XX

PA (SAGR-) SAGRES DISCOVERY.

XX

PI Morris DW, Engelhard EK;

XX

DR WPI; 2003-239337/23.

DR PC:NCBI; gi36552.

DR PC_ENCPRO:NCBI; gi36553.

XX

PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,

PT cancers, neoplasm, adenocarcinoma, or sarcomas.

XX

PS Claim 1; SEQ ID NO 101; 2304pp; English.

XX

CC The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined
 CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a human mRNA of the invention.

CC

CC Revised record issued on 11-JUN-2007 : Enhanced with precomputed

CC information from BOND.

XX

Db	901	CGAGTGGCGGGCGGCCAACTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Qy	961	CGGGCGGCGCGGGCGGTGGGGTTAGCAAACCGCAGCCAAAGCTCATCTTGGCAGGCGGCG	1020
Db	961	CGGGCGGCGCGGGCGGTGGGGTTAGCAAACCGCAGCCAAAGCTCATCTTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCCGCCCTCTCTCGCGCCGAACAGGCGGGG	1080
Db	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCCGCCCTCTCTCGCGCCGAACAGGCGGGG	1080
Qy	1081	CCGCCGCCCTGCTGCCCTTGGGCGCGCGCCGACCACTCGCTGTACAAGGCGCGGA	1140
Db	1081	CCGCCGCCCTGCTGCCCTTGGGCGCGCGCCGACCACTCGCTGTACAAGGCGCGGA	1140
Qy	1141	CTCCAGCGCCTCGGCCCTCCGCCCTCTCGGCAGCCTCGGCCCTCCGAGCGCTCGCGGGCC	1200
Db	1141	CTCCAGCGCCTCGGCCCTCCGCCCTCTCGGCAGCCTCGGCCCTCCGAGCGCTCGCGGGCC	1200
Qy	1201	CGGGCAAGCACCTGGCGGAGAAGAAGTGAAGCGCTCTACCTGTTTGGCGGCCCTGGGCA	1260
Db	1201	CGGGCAAGCACCTGGCGGAGAAGAAGTGAAGCGCTCTACCTGTTTGGCGGCCCTGGGCA	1260
Qy	1261	CGTCGTCGTCGCCCTGGGCGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Db	1261	CGTCGTCGTCGCCCTGGGCGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCCGACGCGCCAGCCTGAGCGGGCGAGCA	1380
Db	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCCGACGCGCCAGCCTGAGCGGGCGAGCA	1380
Qy	1381	GCGCGGCTCGTCCCGCGCGCGGCGCTCGCCCGCGACACCGCGGCTACGCGAGCC	1440
Db	1381	GCGCGGCTCGTCCCGCGCGCGGCGCTCGCCCGCGACACCGCGGCTACGCGAGCC	1440
Qy	1441	TGCGCGCGCCTCGCCCGCCCGTCCAGCGCGCCTCGCACGCTCTCTCTCGGCCCTCGT	1500
Db	1441	TGCGCGCGCCTCGCCCGCCCGTCCAGCGCGCCTCGCACGCTCTCTCTCGGCCCTCGT	1500
Qy	1501	CCCACTCCTCCTCTCTCTCTCTCTCGGGCTCTCGTCTCCGACGACGAGTTTCAAGACG	1560
Db	1501	CCCACTCCTCCTCTCTCTCTCTCTCTCGGGCTCTCGTCTCCGACGACGAGTTTCAAGACG	1560
Qy	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGAGCTTCAGTT	1620
Db	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGAGCTTCAGTT	1620
Qy	1621	CGTCGTCGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTCGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCCGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCCGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACCTACTGAAGGCGCGCAGGCAGGAGAAAGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACCTACTGAAGGCGCGCAGGCAGGAGAAAGGC	1800
Qy	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAGGAAAAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Db	1861	GAGTTTAAAGAGAAAGGAAAAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920

Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTGAGACCCGCGCTCCCATCCCCACCT	1980
Db	1921		1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Db	1981		2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAAGCGAC	2100
Db	2041		2100
Qy	2101	TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101		2160
Qy	2161	TTGCTGCAACCCAGCAAGAAGCGAGTTAGTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161		2220
Qy	2221	CTTCCTGCATCACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Db	2221		2280
Qy	2281	GAACCCAGCGCACCCCTCCCCCCTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281		2340
Qy	2341	GGTGACGAATTGTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Db	2341		2400
Qy	2401	GCAGGGCAATTCCCGTTTGGGGCCTTTTTTCTCCTCCTTTTCCCTTGCCCCCTCTGC	2460
Db	2401		2460
Qy	2461	AGCCGGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCCGGCGCTAGGAAATG	2520
Db	2461		2520
Qy	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Db	2521		2580
Qy	2581	CGAACTGGAAGGGGTTACCGGTCAAACCTGAAATGGATTGACGTTGGGGAGCTGGCGG	2640
Db	2581		2640
Qy	2641	CGGCGGTGCTGGGCCCTCCGCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA	2700
Db	2641		2700
Qy	2701	GACCCCGGAGGCGTGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG	2760
Db	2701		2760
Qy	2761	GGCGAGTGGTTTCGGAAGAAAAAGAAAAAGGG	2797
Db	2761		2797

RESULT 5

ADA02535

ID ADA02535 standard; cDNA; 2797 BP.

XX

AC ADA02535;

XX

DT 11-JUN-2007 (revised)
 DT 06-NOV-2003 (first entry)
 XX
 DE Human SOX4 carcinoma associated cDNA, SEQ ID NO:1053.
 XX
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003057146-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 26-DEC-2002; 2002WO-US041414.
 XX
 PR 26-DEC-2001; 2001US-00035832.
 XX
 PA (SAGR-) SAGRES DISCOVERY.
 XX
 PI Morris DW;
 XX
 DR WPI; 2003-587068/55.
 DR PC:NCBI; gi36552.
 DR PC_ENCPRO:NCBI; gi36553.
 XX
 PT New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 XX
 PS Claim 1; SEQ ID NO 1053; 245pp; English.
 XX
 CC The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed human CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Revised record issued on 11-JUN-2007 : Enhanced with precomputed
 CC information from BOND.
 XX
 SQ Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 9; Length 2797;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTCTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTCTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60

Qy	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAATAAATCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Db	121	GGAATAAATCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGCCAGTTCGGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGCCAGTTCGGCCGCCGCGCG	240
Qy	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCG	300
Db	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCTGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCTGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAAGCGTGCTGGCCGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAAGCGTGCTGGCCGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGCTCCACCGCTCCACGGGCG	480
Db	421	CCGGCTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGCTCCACCGCTCCACGGGCG	480
Qy	481	GCAAGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTATAGTGTGCTGCGAGATCGAGCGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCCTTATAGTGTGCTGCGAGATCGAGCGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAGACAGCG	660
Db	601	TGCACAACGCCGAGATCTCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAGACAGCG	660
Qy	661	ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGCTGACTACC	720
Db	661	ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGCTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781	CGGCCGCCGCTCCTCAAGCCGGGGGAGAAGGAGACAAGTCGGTGGCAGTGGCGGGG	840
Db	781	CGGCCGCCGCTCCTCAAGCCGGGGGAGAAGGAGACAAGTCGGTGGCAGTGGCGGGG	840
Qy	841	GCGGCCATGGGGCGGCGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGCGGTG	900
Db	841	GCGGCCATGGGGCGGCGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGCGGTG	900
Qy	901	CGAGTGGCGCGGCGCCAACTCCAACCGCGCGAGAAAAGAGCTGCGGCTCCAAGTGG	960
Db	901	CGAGTGGCGCGGCGCCAACTCCAACCGCGCGAGAAAAGAGCTGCGGCTCCAAGTGG	960
Qy	961	CGGGCGGCGCGGCGGTGGGGTTAGCAAAACCGCAGCGCAAGCTCATCTGGCAGGCGGCG	1020
Db	961	CGGGCGGCGCGGCGGTGGGGTTAGCAAAACCGCAGCGCAAGCTCATCTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAAGCAGCGGCTGCGCGCGCGCTCCTTCGCGCGGCAACAGCGGGGG	1080

Db	1021	GCGCGGCGGGGAAAGCAGCGGCTGCCGCCGCCCTCCTTCGCCGCCGAACAGCGGGG	1080
Qy	1081	CCGCGCGCCTGCTGCCCTGGGCGCCGCCGACCACTCGTGTACAAGCGCGGA	1140
Db	1081		1140
Qy	1141	CTCCAGCGCCTCGGCCCTCCGCTCCTCGGCAGCCTCGGCCCTCCGAGCGCTCGCGGGC	1200
Db	1141	CTCCAGCGCCTCGGCCCTCCGCTCCTCGGCAGCCTCGGCCCTCCGAGCGCTCGCGGGC	1200
Qy	1201	CGGGCAAGCACTTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTGGCGGCCCTGGGCA	1260
Db	1201	CGGGCAAGCACTTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTGGCGGCCCTGGGCA	1260
Qy	1261	CGTCGTCGTCGCCCGTGGCGGCGTGGCGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Db	1261	CGTCGTCGTCGCCCGTGGCGGCGTGGCGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGGCGCGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Db	1321	TGTACGAGGAGGAGGGCGCGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Qy	1381	GCGCGCCTCGTCCCGCGCGCGCGCTCGCCCGCCGACACCGCGGCTACGCGAGCC	1440
Db	1381	GCGCGCCTCGTCCCGCGCGCGCGCTCGCCCGCCGACACCGCGGCTACGCGAGCC	1440
Qy	1441	TGCGCGCGCCTCGCCCGCCCGTCCAGCGCGCCTCGCAGCGCTCCTCTCGGCCCTCGT	1500
Db	1441	TGCGCGCGCCTCGCCCGCCCGTCCAGCGCGCCTCGCAGCGCTCCTCTCGGCCCTCGT	1500
Qy	1501	CCCACTCCTCCTCTTCTCTCTCTCGGGCTCCTCGTCTCCGACGACGAGTTCAAGACG	1560
Db	1501	CCCACTCCTCCTCTTCTCTCTCTCGGGCTCCTCGTCTCCGACGACGAGTTCAAGACG	1560
Qy	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Db	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Qy	1621	CGTCGTCGGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTCGGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCCGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCCGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTACCTACTGAAGGGCGCGCAGGCAGGAGAGGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTACCTACTGAAGGGCGCGCAGGCAGGAGAGGGC	1800
Qy	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAGGAAAAAGAAAGAAAGTAAGCAGGGCTCGTTGCGCCCGCT	1920
Db	1861	GAGTTTAAAGAGAAAGGAAAAAGAAAGAAAGTAAGCAGGGCTCGTTGCGCCCGCT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTGGACCCGCGTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTGGACCCGCGTCCCATCCCCACCT	1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGAGCGCGGAGGAGAGAGGGTAGAC	2040
Db	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGAGCGCGGAGGAGAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGATGGCAAAAAAAAAAGCGAC	2100

```

Db      2041  AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC  2100
      |
Qy      2101  TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC  2160
      |
Db      2101  TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC  2160
      |
Qy      2161  TTGCTGCACCCCGCAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC  2220
      |
Db      2161  TTGCTGCACCCCGCAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC  2220
      |
Qy      2221  CTTCTGCATCACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA  2280
      |
Db      2221  CTTCTGCATCACCACCTTGGTTTTGTTTTATTTGCTTCTTGGTCAAGAAAGGAGGGGA  2280
      |
Qy      2281  GAAACCGAGCGCACCCCTCCCCCCTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG  2340
      |
Db      2281  GAAACCGAGCGCACCCCTCCCCCCTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG  2340
      |
Qy      2341  GGTGACGAATTTGGCCGATGGCAGATGTTTTGGGGGAACGCCGGGACTGAGAGACTCCAC  2400
      |
Db      2341  GGTGACGAATTTGGCCGATGGCAGATGTTTTGGGGGAACGCCGGGACTGAGAGACTCCAC  2400
      |
Qy      2401  GCAGGCGAATTCCTGTTGGGGCCTTTTTTCTCCCTCTTTTCCCTTGCCCCCTCTGC  2460
      |
Db      2401  GCAGGCGAATTCCTGTTGGGGCCTTTTTTCTCCCTCTTTTCCCTTGCCCCCTCTGC  2460
      |
Qy      2461  AGCCGGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG  2520
      |
Db      2461  AGCCGGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG  2520
      |
Qy      2521  ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGGCGGAGGAGACA  2580
      |
Db      2521  ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGGCGGAGGAGACA  2580
      |
Qy      2581  CGAACTGGAAGGGGGTTACCGTCAAACGAAATGGATTGACAGTTGGGGAGCTGGCGG  2640
      |
Db      2581  CGAACTGGAAGGGGGTTACCGTCAAACGAAATGGATTGACAGTTGGGGAGCTGGCGG  2640
      |
Qy      2641  CGGCGGCTGCTGGGCCCTCCGCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA  2700
      |
Db      2641  CGGCGGCTGCTGGGCCCTCCGCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA  2700
      |
Qy      2701  GACCCCGGAGGCGTGAGGAGAGGAGACTGTTTGTATGTGGTACAGGGGCAGTCAGTGGAG  2760
      |
Db      2701  GACCCCGGAGGCGTGAGGAGAGGAGACTGTTTGTATGTGGTACAGGGGCAGTCAGTGGAG  2760
      |
Qy      2761  GCGGAGTGGTTTCGGAAGAAAAAGAAAAAGGG  2797
      |
Db      2761  GCGGAGTGGTTTCGGAAGAAAAAGAAAAAGGG  2797

```

RESULT 6

ADH28813

ID ADH28813 standard; DNA; 2797 BP.

XX

AC ADH28813;

XX

DT 11-JUN-2007 (revised)

DT 11-MAR-2004 (first entry)

XX

DE Human chronic myelogenous leukaemia (CML) gene marker #81.

XX

KW ds; chronic phase chronic myelogenous leukaemia; CP-CML;

KW blast crisis CML; BC-CML; human; chronic myelogenous leukaemia;

KW gene marker.

XX

```

OS      Homo sapiens.
XX
PN      US2003104426-A1.
XX
PD      05-JUN-2003.
XX
PF      14-JUN-2002; 2002US-00171581.
XX
PR      18-JUN-2001; 2001US-0298914P.
XX
PA      (LINS/) LINSLEY P S.
PA      (MAOM/) MAO M.
PA      (DAIH/) DAI H.
PA      (HEY/) HE Y.
PA      (RAD/) RADICH J P.
XX
PI      Linsley PS, Mao M, Dai H, He Y, Radich JP;
XX
DR      WPI; 2003-787046/74.
DR      PC:NCBI; gi36552.
DR      PC_ENCPRO:NCBI; gi36553.
XX
PT      Classifying cell sample as chronic phase chronic myelogenous leukemia or
PT      blast crisis chronic myelogenous leukemia by detecting difference in
PT      expression of genes corresponding to the markers such as X15415, U89436.
XX
PS      Disclosure; SEQ ID NO 81; 3lpp; English.
XX
CC      The invention relates to a method of classifying a cell sample as chronic
CC      phase chronic myelogenous leukaemia (CP-CML) or blast crisis CML (BC-
CC      CML). The method is useful for classifying a sample as CP-CML or BC-CML.
CC      The present sequence represents a human chronic myelogenous leukaemia
CC      (CML) gene marker used to distinguish blast crisis CML from chronic phase
CC      CML.
CC
CC      Revised record issued on 11-JUN-2007 : Enhanced with precomputed
XX      information from BOND.
XX
SQ      Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match      100.0%; Score 2797; DB 10; Length 2797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TTCCCCAGCATTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1  TTCCCCAGCATTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG 60

Qy      61  ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGACTCTCCAGCCTG 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61  ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGACTCTCCAGCCTG 120

Qy      121  GGAActATAACTCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121  GGAActATAACTCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT 180

Qy      181  TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTCCGCCCGCCGCGCG 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181  TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTCCGCCCGCCGCGCG 240

Qy      241  CGTCTTCCCGTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCGG 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      241  CGTCTTCCCGTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGCGCATCGCGCGG 300

Qy      301  CGGCCCGCCGAGGGTGTGAGCGCGCGTGGGCGCCCGCCGAGCCGAGGCCATGGTGCAGC 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      301  CGGCCCGCCGAGGGTGTGAGCGCGCGTGGGCGCCCGCCGAGCCGAGGCCATGGTGCAGC 360
    
```

Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCCTCGAGTCGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Db	421	CCGGCCTCGAGTCGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTCATGTTGGTTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCCTTCATGTTGGTTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTCTCTCAAGACAGCG	660
Db	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTCTCTCAAGACAGCG	660
Qy	661	ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCTCAAGCACATGGTGACTION	720
Db	661	ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCTCAAGCACATGGTGACTION	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAAGTCCAGCTCCT	780
Db	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAAGTCCAGCTCCT	780
Qy	781	CGGCCCGCGCTCCTTCCAAGCCGGGGAGAAGGGAGACAAGTTCGGTCAGTGGCGGGG	840
Db	781	CGGCCCGCGCTCCTTCCAAGCCGGGGAGAAGGGAGACAAGTTCGGTCAGTGGCGGGG	840
Qy	841	CGGCCATGGGGCGGCGCGCGGCGGCGGAGCAGCAACCGGGGGAGGAGGCGCGGTG	900
Db	841	CGGCCATGGGGCGGCGGCGGCGGCGGAGCAGCAACCGGGGGAGGAGGCGCGGTG	900
Qy	901	CGAGTGGCGGCGGCCAAGTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Db	901	CGAGTGGCGGCGGCCAAGTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Qy	961	CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTTGGCAGGCGGCG	1020
Db	961	CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCCTCCTTCGCGCGCGAACAGGCGGGG	1080
Db	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCCTCCTTCGCGCGCGAACAGGCGGGG	1080
Qy	1081	CCGCGCGCTGCTGCCCTGGGCGCGCGCGCCACCACTCGCTGTACAAGGCGCGGA	1140
Db	1081	CCGCGCGCTGCTGCCCTGGGCGCGCGCGCCACCACTCGCTGTACAAGGCGCGGA	1140
Qy	1141	CTCCAGCGCTCGGCTCCGCTCCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Db	1141	CTCCAGCGCTCGGCTCCGCTCCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Qy	1201	CGGGCAAGCACTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTGGCGGCTGGGCA	1260
Db	1201	CGGGCAAGCACTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTGGCGGCTGGGCA	1260
Qy	1261	CGTCGTCGTCGCGCTGGCGGCGTGGGCGCGGAGCGACCCAGCGACCCCTGGGCC	1320
Db	1261	CGTCGTCGTCGCGCTGGCGGCGTGGGCGCGGAGCGACCCAGCGACCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGCGCGGCTGTCGCGCGACGCGCCAGCTGAGCGGCGCGACGA	1380

Db	1321	TGTACGAGGAGGAGGGCGCGGGCTGCTCGCCCGACGCGCCACGCTGAGCGGCCGACGA	1380
Qy	1381	GCGCCGCCTCGTCCCCCGCGCGCGCGCTCGCCCGCGACACCGCGGCTACGCCAGCC	1440
Db	1381		1440
Qy	1441	TGCGCGCCGCCTCGCCCGCCCCGTCCAGCGCGCCCTCGCACGCGTCTCTCTCGGCCCTCGT	1500
Db	1441		1500
Qy	1501	CCCACTCTCTCTTCTCTCTCTCTCGGGCTCTCTGCTCTCTCGACGACGAGTTTCAAGACG	1560
Db	1501		1560
Qy	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Db	1561		1620
Qy	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCGGCTCCGGCTCGCACT	1680
Db	1621		1680
Qy	1681	TCGAGTTCCCGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681		1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGGCGCGCAGGCAGGGAGAAGGC	1800
Db	1741		1800
Qy	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801		1860
Qy	1861	GAGTTTAAAGAGAAAGGGAAAAAAGAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Db	1861		1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTTGACCCGCGCTCCCATCCCCACCT	1980
Db	1921		1980
Qy	1981	TCCCGGGCCGGGACCCACTCTGCCAGCCGAGGGACGCGGAGGAGGAAGAGGGTAGAC	2040
Db	1981		2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAAGCGAC	2100
Db	2041		2100
Qy	2101	TTGAGTTTGCTCCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101		2160
Qy	2161	TTGCTGCACCCCCAGCAAGAAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161		2220
Qy	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTTTGTCTCTTGGTCAAGAAAGGAGGGGA	2280
Db	2221		2280
Qy	2281	GAACCCAGCGCACCCCTCCCCCTTTTTTAAACGGGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281		2340
Qy	2341	GGTGACGAATTGGCCGATGCGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400

```

Db      2341  GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC 2400
Qy      2401  GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC 2460
Db      2401  GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC 2460
Qy      2461  AGCCGGAGGAGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG 2520
Db      2461  AGCCGGAGGAGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG 2520
Qy      2521  ACCCGAGAACCCTGTTGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA 2580
Db      2521  ACCCGAGAACCCTGTTGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA 2580
Qy      2581  CGAACTGGAAGGGGTTACCGTCAAACCTGAAATGGATTGACGTTGGGGAGCTGGCGG 2640
Db      2581  CGAACTGGAAGGGGTTACCGTCAAACCTGAAATGGATTGACGTTGGGGAGCTGGCGG 2640
Qy      2641  CGGCGGCTGCTGGGCCTCCGCCTTCTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Db      2641  CGGCGGCTGCTGGGCCTCCGCCTTCTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Qy      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGTATGTTGACAGGGGAGTCACTGGAG 2760
Db      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGTATGTTGACAGGGGAGTCACTGGAG 2760
Qy      2761  GGCGAGTGGTTTCGGAAGAAAAAAGAAAAAAGGG 2797
Db      2761  GGCGAGTGGTTTCGGAAGAAAAAAGAAAAAAGGG 2797

```

RESULT 7

ACN38485

ID ACN38485 standard; cDNA; 2797 BP.

XX

AC ACN38485;

XX

DT 11-JUN-2007 (revised)

DT 18-NOV-2004 (first entry)

XX

DE Tumour-associated antigenic target (TAT) cDNA DNA290785, SEQ ID NO:2032.

XX

KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;

KW tumour; diagnosis; cell proliferative disorder; breast cancer;

KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;

KW central nervous system cancer; bladder cancer; pancreatic cancer;

KW cervical cancer; melanoma; leukaemia; hybridisation probe;

KW chromosome identification; chromosome mapping; gene mapping;

KW gene therapy; cytostatic; gene; ss.

XX

OS Homo sapiens.

XX

PN WQ2004030615-A2.

XX

PD 15-APR-2004.

XX

PF 29-SEP-2003; 2003WO-US028547.

XX

PR 02-OCT-2002; 2002US-0414971P.

XX

PA (GETH) GENENTECH INC.

XX

PI Wu TD, Zhang Z, Zhou Y;

XX

DR WPI; 2004-347921/32.

DR

P-PSDB; ABM80791.

DR PC:NCBI; gi36552.
DR PC_ENCPRO:NCBI; gi36553.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 2032; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
CC
CC Revised record issued on 11-JUN-2007 : Enhanced with precomputed
CC information from BOND.
XX
SQ Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 12; Length 2797;
Best Local Similarity 100.0%; Pred. Loc. 0;
Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTGAGAAACTCCTCTCTACTTTAGCACGGTCTCCAGACTACGCCGAGAG	60
Db	1	TTCCCCAGCATTGAGAAACTCCTCTCTACTTTAGCACGGTCTCCAGACTACGCCGAGAG	60
Qy	61	ACAGCAAATGACGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATGACGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAATATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Db	121	GGAATATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTTGCAGGAGTTGAGGGGCCAGTTCCGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTTGCAGGAGTTGAGGGGCCAGTTCCGCCGCCGCGCG	240
Qy	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCCGGGGAACCGGAGGGGCCCGGCATCGCGCG	300
Db	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCCGGGGAACCGGAGGGGCCCGGCATCGCGCG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGCGCCCCCGGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGCGCCCCCGGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGCCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGCCGAGAGCTCGGACTCGGGCG	420

Qy	421	CCGGCCTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Db	421	CCGGCCTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTCATGGTGTGGTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCCTTCATGGTGTGGTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTGCTCAAGACAGCG	660
Db	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTGCTCAAGACAGCG	660
Qy	661	ACAAGATCCCTTTCATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGTGACTACC	720
Db	661	ACAAGATCCCTTTCATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781	CGGCCGCCCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTGGCAGTGGCGGG	840
Db	781	CGGCCGCCCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTGGCAGTGGCGGG	840
Qy	841	GCGGCCATGGGGCGGCGGCGGCGGCGGAGCAGCAACCGGGGGAGGAGGCGGCGGT	900
Db	841	GCGGCCATGGGGCGGCGGCGGCGGCGGAGCAGCAACCGGGGGAGGAGGCGGCGGT	900
Qy	901	CGAGTGGCGGGCGGCGCAACTCCAACCGCGCAGAAAAAGAGCTGCGGCTCCAAGTGG	960
Db	901	CGAGTGGCGGGCGGCGCAACTCCAACCGCGCAGAAAAAGAGCTGCGGCTCCAAGTGG	960
Qy	961	CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTGGCAGGCGGCG	1020
Db	961	CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCGCTCCTTCGCGCCGAACAGGCGGGG	1080
Db	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCGCTCCTTCGCGCCGAACAGGCGGGG	1080
Qy	1081	CCGCCGCCCTGCTGCCCTGGGCGCGCGCGGCGACCACTCGCTGTACAAGGCGCGGA	1140
Db	1081	CCGCCGCCCTGCTGCCCTGGGCGCGCGCGGCGACCACTCGCTGTACAAGGCGCGGA	1140
Qy	1141	CTCCAGCGCCTCGGCTCCGCCTCCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Db	1141	CTCCAGCGCCTCGGCTCCGCCTCCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Qy	1201	CGGGCAAGCACTGGCGGAGAAGAAGGTGAAGCGCTTACCTGTTGGCGGCTGGGCA	1260
Db	1201	CGGGCAAGCACTGGCGGAGAAGAAGGTGAAGCGCTTACCTGTTGGCGGCTGGGCA	1260
Qy	1261	CGTCGTCTGCGCCGTGGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGC	1320
Db	1261	CGTCGTCTGCGCCGTGGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGC	1320
Qy	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGACGA	1380
Db	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGACGA	1380
Qy	1381	GCGCGGCTCGTCGCCCGCGCGGCGCTCGCCCGCGGACACCGGCGCTACGCCAGCC	1440

Db	1381	GCGCCGCCCTCGTCCCCCGCCGCGCGCTCGCCCGCCGACCAACGCGGCTACGCCAGCC	1440
Qy	1441	TGCGCGCCGCCCTCGCCCGCCCGTCCAGCGCGCCCTCGCACGCGTCTCTCGGCCCTCGT	1500
Db	1441	TGCGCGCGCCCTCGCCCGCCCGTCCAGCGCGCCCTCGCACGCGTCTCTCTCGGCCCTCGT	1500
Qy	1501	CCCACTCTCTCTCTCTCTCTCTCTCGGGCTCTCTCGTCTCCGACGACGAGTTCGAAGACG	1560
Db	1501	CCCACTCTCTCTCTCTCTCTCTCTCGGGCTCTCTCGTCTCCGACGACGAGTTCGAAGACG	1560
Qy	1561	ACCTGTCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Db	1561	ACCTGTCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Qy	1621	CGTCGTGCGGCTCGACCGGGACCTGGATTTTAACTTCGAGCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTGCGGCTCGACCGGGACCTGGATTTTAACTTCGAGCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCTCCGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCTCCGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGCGCGCAGGCAGGGAGAAGGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGCGCGCAGGCAGGGAGAAGGGC	1800
Qy	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAAGGAAAAAAGAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCT	1920
Db	1861	GAGTTTAAAGAGAAAAAGGAAAAAAGAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGCTTTTGACCCCGCTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGCTTTTGACCCCGCTCCCATCCCCACCT	1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Db	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAAGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAAGCGAC	2100
Qy	2101	TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101	TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Qy	2161	TTGCTGCACCCCCAGCAAGAAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161	TTGCTGCACCCCCAGCAAGAAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Qy	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTGTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Db	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTGTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Qy	2281	GAAACCCAGCGCACCCCTCCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAAACCCAGCGCACCCCTCCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Db	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Qy	2401	GCAGGCGAATTCCCGTTTGGGCGCTTTTTTTCCTCCCTCTTTTCCCTTGCCCTCTGC	2460

```

Db      2401 GCAGCGCAATTCCCGTTGGGGCCTTTTTTCCTCCCTCTTTTCCCCCTGCCCCCTCTGC 2460
Qy      2461 AGCCGGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG 2520
Db      2461 AGCCGGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG 2520
Qy      2521 ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGCGGAGGAGACA 2580
Db      2521 ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGCGGAGGAGACA 2580
Qy      2581 CGAACTGGAAGGGGTTTCACGGTCAAACCTGAAATGGATTTCACGTTGGGGAGCTGGCGG 2640
Db      2581 CGAACTGGAAGGGGTTTCACGGTCAAACCTGAAATGGATTTCACGTTGGGGAGCTGGCGG 2640
Qy      2641 CGGCGGCTGCTGGGCCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Db      2641 CGGCGGCTGCTGGGCCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Qy      2701 GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
Db      2701 GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
Qy      2761 GGCAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797
Db      2761 GGCAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797

```

RESULT 8

ADW47984

ID ADW47984 standard; cDNA; 2797 BP.

XX

AC ADW47984;

XX

DT 11-JUN-2007 (revised)

DT 07-APR-2005 (first entry)

XX

DE Human sex determining region Y box 4 (SOX4) nucleic acid.

XX

KW Sex determining region Y box 4; SOX4; prostate tumor; cytostatic;
gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 351..1775

FT

FT /*tag= a

FT

FT /product= "Human SOX4"

XX

PN WO2005007830-A2.

XX

PD 27-JAN-2005.

XX

PF 14-JUL-2004; 2004WO-US022850.

XX

PR 14-JUL-2003; 2003US-0487553P.

XX

PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

XX

PI Vanaja DK, Young CYF;

XX

DR WPI; 2005-102097/11.

DR

DR P-PSDB; ADW47985.

DR

DR PC:NCBI; gi36552.

DR

DR PC_ENCPRO:NCBI; gi36553.

XX

PT Detecting/detecting and distinguishing between or among prostate cell

PT proliferative disorders or their stages in a subject, useful for treating
PT prostate cancer, by determining gene expression level of e.g. supervillin
PT (SVIL).
XX
PS Claim 1; SEQ ID NO 43; 178pp; English.
XX
CC The invention provides novel methods and compositions for the diagnosis,
CC staging and prognosis of prostate cancer based on DNA methylation and/or
CC modulation of gene expression, including transcriptional silencing. Gene
CC expression profiling in benign and untreated human prostate cancer
CC tissues identified diagnostic and/or prognostic nucleic acid and protein
CC markers. These included: the differentially (relative to benign tissue)
CC down-regulated sequences corresponding to zinc finger protein 185
CC (ZNF185), prostate secretory protein (PSP94), bulbous pemphigoid antigen
CC (BPAG), supervillin (SVIL), proline rich membrane anchor 1 (PRIMA1),
CC TU3A, FLJ14084, KIAA 1210, Sorbin and SH3 domain containing 1 (SORBS1),
CC and C21orf63; and the differentially up-regulated sequences MARCKS-like
CC protein (MLP), SRY (sex determining region Y)-box 4 (SOX4), fatty acid
CC binding protein 5 (FABP5), MAL2 and Erg isoform 2 (erg-2). Also provided
CC are nucleic acids, nucleic acid arrays and kits useful for detecting, or
CC for detecting and differentiating between or among prostate cell
CC proliferative disorders and/or tumor progression. The present sequence is
CC that of a polynucleotide encoding human SOX4, a high mobility group box 4
CC transcription factor involved in the regulation of embryonic development
CC and in the determination of cell fate. Increase of SOX4 levels indicates
CC a role in development and/or progression of prostate cancer.
CC
CC Revised record issued on 11-JUN-2007 : Enhanced with precomputed
CC information from BOND.
XX
SQ Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 16; Length 2797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTGAGAACTCCTCTCTACTTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTGAGAACTCCTCTCTACTTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Qy	61	ACAGCAAATGCGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATGCGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAACATAAATCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT	180
Db	121	GGAACATAAATCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTCCGGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCGAGTTCCGGCCGCCGCGCG	240
Qy	241	CGTCTTCCCGTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Db	241	CGTCTTCCCGTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCTGGGCGCCCGCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCTGGGCGCCCGCGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCCTCGAGCTGGGAATCGCCTCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Db	421	CCGGCCTCGAGCTGGGAATCGCCTCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480

Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCGCTTCATGGTGTGGTTCGAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCGCTTCATGGTGTGGTTCGAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAAACGCTGGAAGTGCTCAAAGACAGCG	660
Db	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAAACGCTGGAAGTGCTCAAAGACAGCG	660
Qy	661	ACAAGATCCCTTTTCATTTCGAGAGGCGGAGCGGCTGCGGCTCAAGCACATGGTGACTACC	720
Db	661	ACAAGATCCCTTTTCATTTCGAGAGGCGGAGCGGCTGCGGCTCAAGCACATGGTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721	CCGACTACAAGTACCGGCCAGGAAGAAGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781	CGGCCGCGCCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTGGCAGTGGCGGGG	840
Db	781	CGGCCGCGCCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTGGCAGTGGCGGGG	840
Qy	841	GCGGCCATGGGGGCGGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGGCGTG	900
Db	841	GCGGCCATGGGGGCGGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGGCGTG	900
Qy	901	CGAGTGGCGGCGGCCCAACTCCAAACCGGCGCAGAAAAGAGCTGCGGCTCCAAAGTGG	960
Db	901	CGAGTGGCGGCGGCCCAACTCCAAACCGGCGCAGAAAAGAGCTGCGGCTCCAAAGTGG	960
Qy	961	CGGGCGGCGGCGGCGGTGGGGTTAGCAAAACCGCAGCCAAAGTCATCTGGCAGGCGCG	1020
Db	961	CGGGCGGCGGCGGCGGTGGGGTTAGCAAAACCGCAGCCAAAGTCATCTGGCAGGCGCG	1020
Qy	1021	GCGGCGGCGGGAAGCAGCGGCTGCCGCGCGCCTCCTTCGCGCGCGAACAGGCGGGG	1080
Db	1021	GCGGCGGCGGGAAGCAGCGGCTGCCGCGCGCCTCCTTCGCGCGCGAACAGGCGGGG	1080
Qy	1081	CCGCGCGCCTGCTGCCCTTGGGCGCGCGCGGACCACTCGCTGTACAAGGCGCGGA	1140
Db	1081	CCGCGCGCCTGCTGCCCTTGGGCGCGCGCGGACCACTCGCTGTACAAGGCGCGGA	1140
Qy	1141	CTCCAGCGCCTCGGCTCCGCTCCTTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Db	1141	CTCCAGCGCCTCGGCTCCGCTCCTTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Qy	1201	CGGGCAAGCACCTGGCGGAGAAGAAGTGAAGCGGCTACCTGTTGGCGGCCTGGGCA	1260
Db	1201	CGGGCAAGCACCTGGCGGAGAAGAAGTGAAGCGGCTACCTGTTGGCGGCCTGGGCA	1260
Qy	1261	CGTCGTCGTCGCCCTGGGCGGCGTGGGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Db	1261	CGTCGTCGTCGCCCTGGGCGGCGTGGGCGGGAGCCGACCCAGCGACCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCGCGAGCA	1380
Db	1321	TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCGCGAGCA	1380
Qy	1381	GCGCGGCTCGTCCCCGCGCGGCGGCTCGCCGCGGACACCGGGGTACGCCAGCC	1440
Db	1381	GCGCGGCTCGTCCCCGCGCGGCGGCTCGCCGCGGACACCGGGGTACGCCAGCC	1440
Qy	1441	TGCGCGCGCCTCGCCGCGCCGCTCCAGCGCGCCTCGCACGCGCTCCTTCGGCCTCGT	1500

Db	1441	TGCGCGCGCGCTCGCGCGCGCCCTCCAGCGCGCCTCGCACGCGTCTCTCGGCGCTCGT	1500
Qy	1501	CCCACTCTCTCTCTCTCTCTCTCGGGCTCTCGTCTCCGACGACGAGTTTCGAAGACG	1560
Db	1501	CCCACTCTCTCTCTCTCTCTCTCGGGCTCTCGTCTCCGACGACGAGTTTCGAAGACG	1560
Qy	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Db	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCAGCTTCAGTT	1620
Qy	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCCGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCCGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTACCTACTGAAGGGCGCGCAGGCAGGGAGAAGGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTACCTACTGAAGGGCGCGCAGGCAGGGAGAAGGGC	1800
Qy	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Db	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAGAAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTTGACCCCGCGTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGTTTTTGACCCCGCGTCCCATCCCCACCT	1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Db	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATGATGTTGTTGTTGATGGCAAAAAAAGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATGATGTTGTTGTTGATGGCAAAAAAAGCGAC	2100
Qy	2101	TTGAGTTTGCTCCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101	TTGAGTTTGCTCCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCCGGAC	2160
Qy	2161	TTGCTGCAACCCAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161	TTGCTGCAACCCAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Qy	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTTTGTCTTCTTGGTCAAGAAAGGAGGGGA	2280
Db	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTTGCTCTTCTTGGTCAAGAAAGGAGGGGA	2280
Qy	2281	GAAACCCAGCGCACCCCTCCCCCCTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAAACCCAGCGCACCCCTCCCCCCTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTTGGCCGATGGCAGATGTTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Db	2341	GGTGACGAATTTGGCCGATGGCAGATGTTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Qy	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Db	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Qy	2461	AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCAGTGTGACCGGCGCTAGGAAATG	2520

```

Db      2461  AGCCGGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG 2520
Qy      2521  ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGGCGGAGGAGACA 2580
Db      2521  ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGGCGGGGGCGGAGGAGACA 2580
Qy      2581  CGAACTGGAAGGGGTTTCACGGTCAAACCTGAAATGGATTTCACGTTGGGGAGCTGGCGG 2640
Db      2581  CGAACTGGAAGGGGTTTCACGGTCAAACCTGAAATGGATTTCACGTTGGGGAGCTGGCGG 2640
Qy      2641  CGGCGGCTGCTGGGCCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Db      2641  CGGCGGCTGCTGGGCCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
Qy      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
Db      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
Qy      2761  GGCGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797
Db      2761  GGCGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797

```

RESULT 9

AEK60054

ID AEK60054 standard; DNA; 2797 BP.

XX

AC AEK60054;

XX

DT 11-JUN-2007 (revised)

DT 16-NOV-2006 (first entry)

XX

DE Human SOX4 mRNA sequence, SEQ ID NO:41.

XX

KW Diagnostic; therapeutic; cancer; neoplasm; carcinoma; lymphoma;
 hematological disease; immune disorder; breast tumor; endocrine disease;
 gynecology and obstetrics; ds; SOX4.

XX

OS Homo sapiens.

XX

PN US2006204982-A1.

XX

PD 14-SEP-2006.

XX

PF 12-JAN-2006; 2006US-00330726.

XX

PR 22-DEC-2000; 2000US-00747377.

PR 02-MAR-2001; 2001US-00798586.

PR 08-NOV-2001; 2001US-00052482.

XX

PA (MORR/) MORRIS D W.

PA (ENGE/) ENGELHARD E K.

XX

PI Morris DW, Engelhard EK;

XX

DR WPI; 2006-688299/71.

DR PC:NCBI; gi36552.

DR PC_ENCPRO:NCBI; gi36553.

XX

PT Screening for anticancer activity comprises detecting a difference
 between the levels of an expression product of a cancer associated gene
 in a cell in the presence and absence of an anticancer drug candidate.

XX

PS Disclosure; SEQ ID NO 41; 28pp; English.

XX

CC The present invention relates to novel sequences for use in diagnosis and

CC treatment of carcinomas, especially lymphoma carcinomas. In addition, the
 CC present invention describes the use of novel compositions for use in
 CC screening methods. Disclosed is a method for screening for anticancer
 CC activity by detecting a difference between the levels of an expression
 CC product of a cancer associated (CA) gene in a cell in the presence and
 CC absence of an anticancer drug candidate. A difference of at least 50% in
 CC the levels of the expression product in the presence of the anticancer
 CC drug candidate compared to the levels of the expression product in the
 CC absence of the anticancer drug candidate indicates that the anticancer
 CC drug candidate has anticancer activity. The methods and compositions are
 CC useful for screening for anticancer activity and for diagnosing and
 CC treating cancer, specifically breast cancer. Sequences given in AEK60014-
 CC AEK60253 are that of CA nucleic acid sequences including genomic
 CC sequence, mRNA and coding sequence. Note: The sequence data for this
 CC patent did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

CC Revised record issued on 11-JUN-2007 : Enhanced with precomputed
 CC information from BOND.

XX
 SQ Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 19; Length 2797;
 Best Local Similarity 100.0%; Pred. Loc. 0;
 Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Qy	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAATATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Db	121	GGAATATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTTGCAGGAGTTGAGGGGCCAGTTCGGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTTGCAGGAGTTGAGGGGCCAGTTCGGCCGCCGCGCG	240
Qy	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGAGGGGCCCGCATCGCGCGG	300
Db	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGAGGGGCCCGCATCGCGCGG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCTGGGCGCCCGCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCTGGGCGCCCGCGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCTTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Db	421	CCGGCTTCGAGCTGGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Qy	481	GCAAGGCCGACGACCCGAGCTGGTCAAGACCCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTCAAGACCCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTATGGTGTGGTCGAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCCTTATGGTGTGGTCGAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCAAGCGGCTGGGCAACGCTGGAAGTGCTCAAGACAGCG	660

Db	601	 TGCACACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGCTGCTCAAGACACAGCG	660
Qy	661	ACAAGATCCCTTTTCATTTCGAGAGGCGGAGCGGCTGCCTCAAGCACATGGCTGACTACC	720
Db	661	 ACAAGATCCCTTTTCATTTCGAGAGGCGGAGCGGCTGCCTCAAGCACATGGCTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721	 CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781	CGGCCGCGCCTCTCCAAGCCGGGGAGAAGGAGACAAAGTCGGTGGCAGTGGCGGGG	840
Db	781	 CGGCCGCGCCTCTCTCCAAGCCGGGGAGAAGGAGACAAAGTCGGTGGCAGTGGCGGGG	840
Qy	841	GCGGCCATGGGGGCGGCGGCGGCGGGAGCAGCAACCGGGGGAGGAGGCGGCGGTG	900
Db	841	 GCGGCCATGGGGGCGGCGGCGGCGGGAGCAGCAACCGGGGGAGGAGGCGGCGGTG	900
Qy	901	CGAGTGGCGGCGGCCAACTCCAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAGTGG	960
Db	901	 CGAGTGGCGGCGGCCAACTCCAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAGTGG	960
Qy	961	CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCGCAAGCTCATCTGGCAGGCGGCG	1020
Db	961	 CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCGCAAGCTCATCTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCCTCTCTCGCCGCCGAACAGGCGGGG	1080
Db	1021	 GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCCTCTCTCGCCGCCGAACAGGCGGGG	1080
Qy	1081	CCGCGCGCCTGCTGCCCTTGGGCGCGCGCGGACCACTCGTGTACAAAGCGCGGA	1140
Db	1081	 CCGCGCGCCTGCTGCCCTTGGGCGCGCGCGGACCACTCGTGTACAAAGCGCGGA	1140
Qy	1141	CTCCAGCGCCTCGGCTCCGCTCTCGGCAGCTTCGGCTTCGCGAGCGCTCGCGGCC	1200
Db	1141	 CTCCAGCGCCTCGGCTCCGCTCTCGGCAGCTTCGGCTTCGCGAGCGCTCGCGGCC	1200
Qy	1201	CGGGCAAGCACCTGGCGGAGAAGAAGGTGAAGCGCTGTACCTGTTGGCGGCTGGGCA	1260
Db	1201	 CGGGCAAGCACCTGGCGGAGAAGAAGGTGAAGCGCTGTACCTGTTGGCGGCTGGGCA	1260
Qy	1261	CGTGTGTGCGCCGTGGGCGGCGTGGGCGCGGAGCCGACCCAGCGACCCCTGGGCC	1320
Db	1261	 CGTGTGTGCGCCGTGGGCGGCGTGGGCGCGGAGCCGACCCAGCGACCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGCGCGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Db	1321	 TGTACGAGGAGGAGGCGCGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Qy	1381	GCGCCGCTCGTCCCGCGCGCGGCGCTCGCCCGCCGACCAACGCGGCTACGCCAGCC	1440
Db	1381	 GCGCCGCTCGTCCCGCGCGCGGCGCTCGCCCGCCGACCAACGCGGCTACGCCAGCC	1440
Qy	1441	TGGCGCGCGCTCGCCGCGCCGTCACGCGCGCTCGACGCGTCTCTCTGGGCTCGT	1500
Db	1441	 TGGCGCGCGCTCGCCGCGCCGTCACGCGCGCTCGACGCGTCTCTCTGGGCTCGT	1500
Qy	1501	CCCACTCTCTCTTCTCTCTCTCGGGCTCTCTGCTCTCCGACGACGAGTTCGAAGACG	1560
Db	1501	 CCCACTCTCTCTTCTCTCTCTCGGGCTCTCTGCTCTCCGACGACGAGTTCGAAGACG	1560
Qy	1561	ACCTGTCTGACCTGAACCCAGCTCAAACCTTGAGAGCATGTCCCTGGGCGAGTTTCAGTT	1620
Db	1561	 ACCTGTCTGACCTGAACCCAGCTCAAACCTTGAGAGCATGTCCCTGGGCGAGTTTCAGTT	1620

Qy	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTCGGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCCGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCCGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGGCGCGCAGGCAGGAGAGAAGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACTACTGAAGGGCGCGCAGGCAGGAGAGAAGGC	1800
Qy	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAGAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAGAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAAGGAAAAAGAAAAAGTAAAGCAGGGCTCGTTTCGCCCGCGT	1920
Db	1861	GAGTTTAAAGAGAAAAAGGAAAAAGAAAAAGTAAAGCAGGGCTCGTTTCGCCCGCGT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGGCGCGTTTTTGACCCCGCGTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGGCGCGTTTTTGACCCCGCGTCCCATCCCCACCT	1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGAGCGCGAGGAGGAAGAGGGTAGAC	2040
Db	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGGAGCGCGAGGAGGAAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Qy	2101	TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101	TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCTTCCAACGAGCTTCCGGAC	2160
Qy	2161	TTGTCTGCACCCAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161	TTGTCTGCACCCAGCAAGAAGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Qy	2221	CTTCTGCATCACCACTTGGTTTTGTTTTATTGTTGCTCTTTGGTCAAGAAAGGAGGGGA	2280
Db	2221	CTTCTGCATCACCACTTGGTTTTGTTTTATTGTTGCTCTTTGGTCAAGAAAGGAGGGGA	2280
Qy	2281	GAAACGAGCGACCCCTCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAAACGAGCGACCCCTCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGATCCAC	2400
Db	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGATCCAC	2400
Qy	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTGTC	2460
Db	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTGTC	2460
Qy	2461	AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG	2520
Db	2461	AGCCGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGGCGCTAGGAAATG	2520
Qy	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Db	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Qy	2581	CGAACTGGAAGGGGTTACCGTCAAACGAAATGGATTGACAGTTGGGGAGCTGGCGG	2640
Db	2581	CGAACTGGAAGGGGTTACCGTCAAACGAAATGGATTGACAGTTGGGGAGCTGGCGG	2640

```

Qy      2641  CGGCGGCTGCTGGGCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCA 2700
Db      2641  CGGCGGCTGCTGGGCCTCCGCCTTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCA 2700
      |||
Qy      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
Db      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGCAGTCAGTGGAG 2760
      |||
Qy      2761  GCGGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797
Db      2761  GCGGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797
      |||

```

RESULT 10

AEM95802

ID AEM95802 standard; cDNA; 2797 BP.

XX

AC AEM95802;

XX

DT 11-JUN-2007 (revised)

DT 22-MAR-2007 (first entry)

XX

DE Human CML marker gene, SEQ ID NO: 81.

XX

KW ss; gene; chronic myelocytic leukemia; genetic marker; diagnostic;

KW DNA microarray; gene expression.

XX

OS Homo sapiens.

XX

PN US2006292623-A1.

XX

PD 28-DEC-2006.

XX

PF 25-AUG-2006; 2006US-00510798.

XX

PR 18-JUN-2001; 2001US-0298914P.

PR 14-JUN-2002; 2002US-00171581.

XX

PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.

PA (ROSE-) ROSETTA INPHARMATICS LLC.

XX

PI Linsley PS, Mao M, Dai H, He Y, Radich JP;

XX

DR WPI; 2007-173869/17.

DR EMBL; X70683.

DR PC:NCBI; gi36552.

DR PC_ENCPRO:NCBI; gi36553.

XX

PT Classification of cell sample as chronic phase or blast crisis of chronic

PT myeloid leukemia involves detecting a difference or similarity in

PT expression by genes from sample corresponding to a set of specific

PT markers, relative to control.

XX

PS Claim 5; SEQ ID NO 81; 28pp; English.

XX

CC The invention relates to classifying a cell sample as chronic phase

CC chronic myeloid leukemia (CP-CML) or blast crisis chronic myeloid

CC leukemia (BC-CML) by detecting a difference in the expression by the cell

CC sample of a first set of several genes relative to a control. The first

CC set of several genes consist of at least 5 genes corresponding to 366

CC markers (SEQ ID Nos: 1 - 366, AEM95722-AEM96087). The invention comprises

CC a kit for determining the progression status of a tissue sample or a

CC sample of bodily fluid from a human or an animal, comprising at least two

CC microarrays, each comprising at least 20 of the 366 markers as given in

CC the specification, and a computer system for determining the similarity

CC of the level of nucleic acid derived from the markers in a sample to that

in the CP-CML template and the BC-CML template. The method involves computing the aggregate differences in expression of each marker between the sample and the CP-CML pool or the BC-CML pool, or determining the correlation of expression, calculated according to the equation as given in the specification, of the markers in the sample to the expression in the CP-CML and BC-CML pools. The microarray for distinguishing CP-CML from BC-CML cell samples comprises a positionally-addressable array of polynucleotide probes bound to a support, the polynucleotide probes comprising a sequence complementary and hybridizable to a different gene corresponding to one of the markers. The method for classifying a sample as CP-CML or BC-CML involves labeling and detecting nucleic acids derived from a sample with a first fluorophore to obtain a first pool of fluorophore-labeled nucleic acids, labeling with a second fluorophore a first pool of nucleic acids derived from two or more CP-CML samples, and a second pool of nucleic acids derived from two or more BP-CML samples, contacting the first fluorophore-labeled nucleic acid and the first pool of second fluorophore-labeled nucleic acid with the first microarray under hybridization conditions, and contacting the first fluorophore-labeled nucleic acid and the second pool of second fluorophore-labeled nucleic acid with the second microarray under hybridization conditions. The similarity between the expression of the markers is calculated by determining a first sum of the differences of expression levels for each marker between the first fluorophore-labeled nucleic acid and the first pool of second fluorophore-labeled nucleic acid, and a second sum of the differences of expression levels for each marker between the first fluorophore-labeled nucleic acid and the second pool of second fluorophore-labeled nucleic acid, where if the first sum is greater than the second sum, the sample is classified as CP-CML, and if the second sum is greater than the first sum, the sample is classified as BC-CML. The method provides accurate determination of different phases (chronic phase or blast crisis) of chronic myeloid leukemia (CML), thereby allowing the determination of treatment options, prognosis and likelihood of therapeutic response. The present sequence is that of one of the human marker genes of the current invention.

Revised record issued on 11-JUN-2007 : Enhanced with precomputed information from BOND.

Sequence 2797 BP; 547 A; 880 C; 931 G; 439 T; 0 U; 0 Other;

Query Match 100.0%; Score 2797; DB 22; Length 2797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTCGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Qy	61	ACAGCAAATCGACGCGGTGAGAGAGCGAGAGAGAGGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATCGACGCGGTGAGAGAGCGAGAGAGAGGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAATATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT	180
Db	121	GGAATATAACTCCTCTGCGAGAGGCGGAGAACTCCTTCCCCAAATCTTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCCAGTTCCGGCCGCCGCGG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCCAGTTCCGGCCGCCGCGG	240
Qy	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Db	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGGCGCCCGCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGGCGCCCGCGAGCCGAGGCCATGGTGCAGC	360

Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCCTCGAGTCGGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Db	421	CCGGCCTCGAGTCGGGAATCGCCTCCTCCCCACGCCCGGCTCCACCGCCTCCACGGGCG	480
Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCTTCATGTTGGTTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCCTTCATGTTGGTTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTCTCTCAAGACAGCG	660
Db	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTCTCTCAAGACAGCG	660
Qy	661	ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCTCAAGCACATGGTGACTACC	720
Db	661	ACAAGATCCCTTTTATTTCGAGAGGCGGAGCGGCTGCGCTCAAGCACATGGTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Db	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780
Qy	781	CGGCCCGCGCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTCAGTGGCGGGG	840
Db	781	CGGCCCGCGCTCCTCCAAGCCGGGGAGAAGGGAGACAAGGTCGGTCAGTGGCGGGG	840
Qy	841	GCGGCCATGGGGGCGGCGGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGGCGTG	900
Db	841	GCGGCCATGGGGGCGGCGGCGGCGGCGGAGCAGCAACGCGGGGGAGGAGGCGGCGTG	900
Qy	901	CGAGTGGCGGCGGCGCAACTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Db	901	CGAGTGGCGGCGGCGCAACTCCAAACCGGCGCAGAAAAAGAGCTGCGGCTCCAAAGTGG	960
Qy	961	CGGGCGGCGGGGCGGTGGGGTTAGCAAACCGCACGCGCAAGCTCATCTTGGCAGGCGGCG	1020
Db	961	CGGGCGGCGGGGCGGTGGGGTTAGCAAACCGCACGCGCAAGCTCATCTTGGCAGGCGGCG	1020
Qy	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCGCTCCTTCGCGCGCGAACAGGCGGGG	1080
Db	1021	GCGGCGGCGGGAAAGCAGCGGCTGCCGCGCGCGCTCCTTCGCGCGCGAACAGGCGGGG	1080
Qy	1081	CCGCGCGCTGCTGCCCTGGGCGCGCGCGCCACCACCTCGCTGTACAAGGCGCGGA	1140
Db	1081	CCGCGCGCTGCTGCCCTGGGCGCGCGCGCCACCACCTCGCTGTACAAGGCGCGGA	1140
Qy	1141	CTCCAGCGCTCGGCTCCGCTCCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Db	1141	CTCCAGCGCTCGGCTCCGCTCCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Qy	1201	CGGGCAAGCACTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTGGCGGCTGGGCA	1260
Db	1201	CGGGCAAGCACTGGCGGAGAAGAAGTGAAGCGCGTCTACCTGTTGGCGGCTGGGCA	1260
Qy	1261	CGTCGTCGTCGCGCTGGCGGCGTGGGCGCGGGAGCGACCCAGCGACCCCTGGGCC	1320
Db	1261	CGTCGTCGTCGCGCTGGCGGCGTGGGCGCGGGAGCGACCCAGCGACCCCTGGGCC	1320
Qy	1321	TGTACGAGGAGGAGGCGCGGGCTGCTCGCCGACGCGGCCAGCTGAGCGGCGCGACGA	1380

Db	1321	TGTACGAGGAGGAGGGCGCGGGCTGCTCGCCCGACGCGCCACGCTGAGCGGCGCAGCA	1380
Qy	1381	GGCGCGCCTCGTCCCCCGCGCGGCGCTCGCCCGCGACACCGCGGCTACGCCAGCC	1440
Db	1381	GGCGCGCCTCGTCCCCCGCGCGGCGCTCGCCCGCGACACCGCGGCTACGCCAGCC	1440
Qy	1441	TGCGCGCGCCTCGCGCGCCCGCTCCAGCGCGCCTCGCACGCTCTCTCTCGGCCTCGT	1500
Db	1441	TGCGCGCGCCTCGCGCGCCCGCTCCAGCGCGCCTCGCACGCTCTCTCTCGGCCTCGT	1500
Qy	1501	CCCACTCTCTCTTCTCTCTCTCTCGGGCTCTCTGCTCTCCGACGACGAGTTTCGAAGACG	1560
Db	1501	CCCACTCTCTCTTCTCTCTCTCTCGGGCTCTCTGCTCTCCGACGACGAGTTTCGAAGACG	1560
Qy	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGCAGCTTCAGTT	1620
Db	1561	ACCTGCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCTGGGCAGCTTCAGTT	1620
Qy	1621	CGTCGTGCGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Db	1621	CGTCGTGCGCGCTCGACCGGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACT	1680
Qy	1681	TCGAGTTCCTGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Db	1681	TCGAGTTCCTGGACTACTGCACGCCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCG	1740
Qy	1741	AGTCACGATCTCCAACCTGGTTTTACCTACTGAAGGGCGCGCAGGCAGGGAGAAGGGC	1800
Db	1741	AGTCACGATCTCCAACCTGGTTTTACCTACTGAAGGGCGCGCAGGCAGGGAGAAGGGC	1800
Qy	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAAGAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Db	1861	GAGTTTAAAGAGAAAAGGAAAAAAGAAAAGAAAAGTAAGCAGGGCTCGTTCGCCCGCGT	1920
Qy	1921	TCTCGTCGTGGATCAAGAGAGCGCGCGCGTTTTTGAGCCCGCTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTGGATCAAGAGAGCGCGCGCGTTTTTGAGCCCGCTCCCATCCCCACCT	1980
Qy	1981	TCCGGGCGCGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Db	1981	TCCGGGCGCGGGACCCACTCTGCCAGCCGAGGGACGCGAGGAGGAAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATGATGTTGTTGTTGATGGCAAAAAAAGAGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATGATGTTGTTGTTGATGGCAAAAAAAGAGCGAC	2100
Qy	2101	TTGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCGGAC	2160
Db	2101	TTGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCCCTTCCAACGAGCTTCGGAC	2160
Qy	2161	TTGCTGCAACCCCGACAAGAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCC	2220
Db	2161	TTGCTGCAACCCCGACAAGAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCC	2220
Qy	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTTTGTCTTCTTGTTCAAGAAAGGAGGGA	2280
Db	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTTTGTCTTCTTGTTCAAGAAAGGAGGGA	2280
Qy	2281	GAACCCAGCGCACCTTCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAACCCAGCGCACCTTCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTTGGCCGATGCGAGATGTTTGGGGGAACGCGGGACTGAGAGACTCCAC	2400

```

Db      2341  GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC 2400
      |||
Qy      2401  GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC 2460
      |||
Db      2401  GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCTCCCTCTTTTCCCTTGCCCCCTCTGC 2460
      |||
Qy      2461  AGCCGAGGAGGAGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGCGCTAGGAAATG 2520
      |||
Db      2461  AGCCGAGGAGGAGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTGACCGCGCTAGGAAATG 2520
      |||
Qy      2521  ACCCGAGAACCCTGTTGGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA 2580
      |||
Db      2521  ACCCGAGAACCCTGTTGGAAGCGCAGCAGCGGAGCTAGGGCGGGGGCGGAGGAGACA 2580
      |||
Qy      2581  CGAACTGGAAGGGGTTACCGTCAAACCTGAAATGGATTGACGTTGGGGAGCTGGCGG 2640
      |||
Db      2581  CGAACTGGAAGGGGTTACCGTCAAACCTGAAATGGATTGACGTTGGGGAGCTGGCGG 2640
      |||
Qy      2641  CGGCGGCTGCTGGGCCTCCGCCTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
      |||
Db      2641  CGGCGGCTGCTGGGCCTCCGCCTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA 2700
      |||
Qy      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGTATGTTGTTACAGGGGCAGTCAGTGGAG 2760
      |||
Db      2701  GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGTATGTTGTTACAGGGGCAGTCAGTGGAG 2760
      |||
Qy      2761  GGCAGTGGTTTCGGAAGAAAAAAGAAAAAAGGG 2797
      |||
Db      2761  GGCAGTGGTTTCGGAAGAAAAAAGAAAAAAGGG 2797

```

RESULT 11

ADE95783

ID ADE95783 standard; DNA; 2797 BP.

XX

AC ADE95783;

XX

DT 12-FEB-2004 (first entry)

XX

DE Human DNA related to SOX4 gene mRNA.

XX

KW cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;

KW lymphoma; breast cancer; prostate cancer; leukaemia; ds; human; SOX4.

XX

OS Homo sapiens.

XX

PN WO2003039484-A2.

XX

PD 15-MAY-2003.

XX

PF 08-NOV-2002; 2002WO-US036071.

XX

PR 08-NOV-2001; 2001US-00052482.

XX

PA (SAGR-) SAGRES DISCOVERY.

XX

PI Morris DW, Engelhard EK;

XX

DR WPI; 2003-441462/41.

XX

PT New carcinoma associated nucleic acids and proteins, useful for screening

PT drug candidates, or for diagnosing and treating carcinomas, e.g.

PT lymphoma, breast cancer, prostate cancer or leukemia.

XX

PS Claim 1; SEQ ID NO 41; 793pp; English.

XX

CC This invention relates to novel recombinant nucleic acids for use in
 CC diagnosis and treatment of cancer, especially carcinomas, as well as the
 CC use of compositions in screening methods. The compositions of the
 CC invention may have cytostatic activity whilst the disclosed sequences may
 CC be useful for gene therapy. The carcinoma associated nucleic acids and
 CC proteins are useful for diagnosing and treating carcinomas, for example
 CC lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
 CC drug candidates or bioactive agents capable of binding to, or modulating
 CC the activity of, a carcinoma associated protein. The present sequence is
 CC a DNA sequence which represents the mRNA derived from the human SOX4 gene
 CC which is a carcinoma associated gene of the invention.

XX
 SQ Sequence 2797 BP; 547 A; 881 C; 930 G; 439 T; 0 U; 0 Other;

Query Match 99.9%; Score 2795.4; DB 9; Length 2797;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2796; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	TTCCCCAGCATTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Db	1	TTCCCCAGCATTGAGAACTCCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAG	60
Qy	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGAGAGAGACTCTCCAGCCTG	120
Db	61	ACAGCAAATGCAGCGCGGTGAGAGAGCGAGAGAGAGGAGAGAGACTCTCCAGCCTG	120
Qy	121	GGAACTATAACTCCTCTGCGAGAGGCGGAGAACTCTTCCCAAATCTTTTGGGGACTTT	180
Db	121	GGAACTATAACTCCTCTGCGAGAGGCGGAGAACTCTTCCCAAATCTTTTGGGGACTTT	180
Qy	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGCCAGTTGCGCCGCCGCGCG	240
Db	181	TCTCTCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGCCAGTTGCGCCGCCGCGCG	240
Qy	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Db	241	CGTCTTCCCGTTTCGGCGTGTGCTTGGCCCGGGGAACCGGGAGGGCCCGGCATCGCGCGG	300
Qy	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Db	301	CGGCCGCCGCGAGGGTGTGAGCGCGCGTGGCGCCGCCGAGCCGAGGCCATGGTGCAGC	360
Qy	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Db	361	AAACCAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCG	420
Qy	421	CCGGCCTCGAGCTGGGAATCGCCTCTCCCCACGCCCGGCTCCACCGCTCCACGGGCG	480
Db	421	CCGGCCTCGAGCTGGGAATCGCCTCTCCCCACGCCCGGCTCCACCGCTCCACGGGCG	480
Qy	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Db	481	GCAAGGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGA	540
Qy	541	ACGCCCTCATGGTGTGGTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Db	541	ACGCCCTCATGGTGTGGTCGCAGATCGAGCGGCGCAAGATCATGGAGCAGTCGCCCGACA	600
Qy	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTCTCTCAAGACACGCG	660
Db	601	TGCACAACGCCGAGATCTCCAAGCGGCTGGGCAACGCTGGAAGTCTCTCAAGACACGCG	660
Qy	661	ACAAGATCCCTTTTATTGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGTGACTACC	720
Db	661	ACAAGATCCCTTTTATTGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGTGACTACC	720
Qy	721	CCGACTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCCT	780

Db	721		CCGACTACAGTACCGGCCAGGAAGAAGGTGAAGTCGGCAACGCCAACTCCAGCTCCT	780
Qy	781		CGGCCGCCCTCTCCAAAGCCGGGGGAGAAGGGAGACAAGGTCGGTGGCAGTGGCGGGG	840
Db	781		CGGCCGCCCTCTCCAAAGCCGGGGGAGAAGGGAGACAAGGTCGGTGGCAGTGGCGGGG	840
Qy	841		GCGGCCATGGGGCGCGCGCGCGGGAGCAGCAACCGGGGGAGGAGGCGCGGTG	900
Db	841		GCGGCCATGGGGCGCGCGCGCGGGAGCAGCAACCGGGGGAGGAGGCGCGGTG	900
Qy	901		CGAGTGGCGGGCGGCCAACTCCAAACCGCGCAGAAAAAGAGCTGCGGCTCCAAGTGG	960
Db	901		CGAGTGGCGGGCGGGCCAACTCCAAACCGCGCAGAAAAAGAGCTGCGGCTCCAAGTGG	960
Qy	961		CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTGGCAGGCGGCG	1020
Db	961		CGGGCGGCGGGCGGTGGGGTTAGCAAACCGCACGCCAAGCTCATCTGGCAGGCGGCG	1020
Qy	1021		GCGGCGGCGGGAAGCAGCGGCTGCCGCCCGCCTCTCTTCCGCCCGAACAGGCGGGG	1080
Db	1021		GCGGCGGCGGGAAGCAGCGGCTGCCGCCCGCCTCTCTTCCGCCCGAACAGGCGGGG	1080
Qy	1081		CCGCCGCCCTGCTGCCCTGGGCGCGCGCGGACCACTCGCTGTACAAGGCGCGGA	1140
Db	1081		CCGCCGCCCTGCTGCCCTGGGCGCGCGCGGACCACTCGCTGTACAAGGCGCGGA	1140
Qy	1141		CTCCAGCGCCTCGGCTCCGCCTCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Db	1141		CTCCAGCGCCTCGGCTCCGCCTCTCGGCAGCTCGGCTCCGAGCGCTCGCGGCC	1200
Qy	1201		CGGGCAAGCACTGGCGGAGAAGGTGAAGCGCGTCTACCTGTTGGCGGCCTGGGCA	1260
Db	1201		CGGGCAAGCACTGGCGGAGAAGGTGAAGCGCGTCTACCTGTTGGCGGCCTGGGCA	1260
Qy	1261		CGTCGTCGTCGCCGCTGGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGC	1320
Db	1261		CGTCGTCGTCGCCGCTGGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGC	1320
Qy	1321		TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Db	1321		TGTACGAGGAGGAGGGCGGGGCTGCTCGCCGACGCGCCAGCCTGAGCGGCCGAGCA	1380
Qy	1381		GCGCCGCTCGTCCCGCGCGCGGCGCTCGCCCGCGGACCAACCGGGCTACGCCAGCC	1440
Db	1381		GCGCCGCTCGTCCCGCGCGCGGCGCTCGCCCGCGGACCAACCGGGCTACGCCAGCC	1440
Qy	1441		TGCGCGCGCTCGCCCGCCCGTCCAGCGCGCTCGCACGCGTCTCTCGGCTCGT	1500
Db	1441		TGCGCGCGCTCGCCCGCCCGTCCAGCGCGCTCGCACGCGTCTCTCGGCTCGT	1500
Qy	1501		CCCACTCTCTCTTCTCTCTCTCGGGCTCTCGTCTCCGACGACGAGTTCAAGACG	1560
Db	1501		CCCACTCTCTCTTCTCTCTCTCGGGCTCTCTCGTCTCCGACGACGAGTTCAAGACG	1560
Qy	1561		ACCTGCTCGACTGAACCCAGCTCAAACCTTGAGAGCATGTCCCTGGGAGCTTCAGTT	1620
Db	1561		ACCTGCTCGACTGAACCCAGCTCAAACCTTGAGAGCATGTCCCTGGGAGCTTCAGTT	1620
Qy	1621		CGTCGTCGCGCTCGACCGGACCTGGATTTAACTTCGAGCCGGCTCCGGCTCGCACT	1680
Db	1621		CGTCGTCGCGCTCGACCGGACCTGGATTTAACTTCGAGCCGGCTCCGGCTCGCACT	1680
Qy	1681		TCGAGTTCGGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGAGACTGGCTCG	1740
Db	1681		TCGAGTTCGGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGAGACTGGCTCG	1740

Qy	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACCTACTGAAGGCGCGCAGGCAGGGAGAAGGGC	1800
Db	1741	AGTCCAGCATCTCCAACCTGGTTTTTCACCTACTGAAGGCGCGCAGGCAGGGAGAAGGGC	1800
Qy	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Db	1801	CGGGGGGGTAGGAGAGGAGAAAAAAAGTGAAAAAAGAAACGAAAGGACAGACGAA	1860
Qy	1861	GAGTTTAAAGAGAAAAGGGAAGAAAAAGAAAAAGTAAGCAGGGCTCGTTGCGCCCGCT	1920
Db	1861	GAGTTTAAAGAGAAAAGGGAAGAAAAAGAAAAAGTAAGCAGGGCTCGTTGCGCCCGCT	1920
Qy	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGCTTTTGACCCGCGCTCCCATCCCCACCT	1980
Db	1921	TCTCGTCGTCGGATCAAGGAGCGCGCGCGCTTTTGACCCGCGCTCCCATCCCCACCT	1980
Qy	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGAGCGGAGGAGGAAGAGGGTAGAC	2040
Db	1981	TCCCGGGCGGGGACCCACTCTGCCAGCCGAGGAGCGGAGGAGGAAGAGGGTAGAC	2040
Qy	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Db	2041	AGGGGCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAGCGAC	2100
Qy	2101	TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCTTCCAACGAGCTTCCGGAC	2160
Db	2101	TTCGAGTTTGCTCCCTTTGCTTGAAGAGACCCCTCCCTTCCAACGAGCTTCCGGAC	2160
Qy	2161	TTGTCTGCACCCAGCAGAAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Db	2161	TTGTCTGCACCCAGCAGAAGGCGAGTTAGTTTTCTAGAGACTTGAAGGAGTCTCCCC	2220
Qy	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Db	2221	CTTCTGCATCACCACCTTGGTTTTGTTTTATTTTGCTTCTTGGTCAAGAAAGGAGGGGA	2280
Qy	2281	GAAACCCAGCGCACCCCTCCCCCTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Db	2281	GAAACCCAGCGCACCCCTCCCCCTTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGG	2340
Qy	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Db	2341	GGTGACGAATTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCAC	2400
Qy	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Db	2401	GCAGGCGAATTCCCGTTTGGGGCCTTTTTTCTCCCTCTTTTCCCTTGCCCCCTCTGC	2460
Qy	2461	AGCCGGAGGAGAGATGTTGAGGGGAGGAGGCCAGGCAGTGTACCGGCGCTAGGAAATG	2520
Db	2461	AGCCGGAGGAGAGATGTTGAGGGGAGGAGGCCAGGCAGTGTACCGGCGCTAGGAAATG	2520
Qy	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Db	2521	ACCCGAGAACCCCGTTGGAAGCGCAGCAGCGGGAGCTAGGGCGGGGGCGGAGGAGACA	2580
Qy	2581	CGAACTGGAAGGGGTTACCGTCAAACCTGAAATGGATTGACGTTGGGGAGCTGGCGG	2640
Db	2581	CGAACTGGAAGGGGTTACCGTCAAACCTGAAATGGATTGACGTTGGGGAGCTGGCGG	2640
Qy	2641	CGGCGGCTGCTGGGCCTCCGCCTTCTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA	2700
Db	2641	CGGCGGCTGCTGGGCCTCCGCCTTCTTTCTACGTGAAATCAGTGAGGTGAGACTTCCCA	2700
Qy	2701	GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGACGTGAGTGGAG	2760
Db	2701	GACCCCGGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGACGTGAGTGGAG	2760

Qy 2761 GGCGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797
 ||||||||||||||||||||||||||||||||||||
 Db 2761 GGCGAGTGGTTTCGGAAGAAAAAGAAAAAGGG 2797

RESULT 12

ADB75556

ID ADB75556 standard; cDNA; 3233 BP.

XX

AC ADB75556;

XX

DT 04-DEC-2003 (first entry)

XX

DE Prostate cancer marker cDNA.

XX

KW Prostate; cancer; cytostatic; gene therapy; marker; ss.

XX

OS Homo sapiens.

XX

PN WO2003009814-A2.

XX

PD 06-FEB-2003.

XX

PF 25-JUL-2002; 2002WO-US023913.

XX

PR 25-JUL-2001; 2001US-0307982P.

PR 22-AUG-2001; 2001US-0314356P.

PR 25-SEP-2001; 2001US-0325020P.

PR 12-DEC-2001; 2001US-0341746P.

PR 05-MAR-2002; 2002US-0362158P.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;

PI Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;

XX

DR WPI; 2003-248033/24.

XX

PT New nucleic acid molecule, useful for diagnosing or treating prostate cancer.

XX

PS Claim 1; SEQ ID NO 380; 99pp; English.

XX

CC The invention relates to newly discovered cancer markers associated with
 CC the cancerous state of prostate cells. Also disclosed is a method of
 CC assessing whether a patient is afflicted with prostate cancer. The method
 CC of the invention involves assessing whether a patient is afflicted with
 CC prostate cancer by comparing the level of expression of a marker in a
 CC patient sample and the normal level of expression of the marker in a
 CC control non-prostate cancer sample, where a significant increase in the
 CC level of expression of the marker in the patient sample and the normal
 CC level indicates that the patient is afflicted with prostate cancer.
 CC Nucleic acids of the invention are useful for diagnosing or treating
 CC prostate cancer, and may be useful in gene therapy. Sequences given in
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 3233 BP; 662 A; 979 C; 1043 G; 547 T; 0 U; 2 Other;

Query Match 99.9%; Score 2793; DB 8; Length 3233;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 CCAGCATTCGAGAACTCTCTCTACTTTAGCACGGTCTCCAGACTCAGCCGAGAGACAG 64

Db	441		CCAGCATTTCGAGAAATCTCTCTACTTTAGACAGGGTCTCCAGACTCAGCCGAGAGACAG	500
Qy	65		CAAACATGCAGCGCGGTGAGAGAGCGAGAGAGAGGAGAGAGACTCTCCAGCCTGGGAA	124
Db	501		CAAACATGCAGCGCGGTGAGAGAGCGAGAGAGAGGAGAGAGACTCTCCAGCCTGGGAA	560
Qy	125		CTATAACTCTCTCTCGAGAGGCGGAGAACTCTTCCCCAAATCTTTTGGGGACTTTTCTC	184
Db	561		CTATAACTCTCTCTCGAGAGGCGGAGAACTCTTCCCCAAATCTTTTGGGGACTTTTCTC	620
Qy	185		TCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCCAGTTCGGCCGCCGCGCGCTC	244
Db	621		TCTTTACCCACCTCCGCCCTGCGAGGAGTTGAGGGGCCAGTTCGGCCGCCGCGCGCTC	680
Qy	245		TTCGCGTTCGGCGTGTGCTTTGGCCCGGGGAACCGGGAGGGCCCGCGGATCGCGCGGCGGC	304
Db	681		TTCGCGTTCGGCGTGTGCTTTGGCCCGGGGAACCGGGAGGGCCCGCGGATCGCGCGGCGGC	740
Qy	305		CGCCGCGAGGGTGTGAGCGCGCGTGGCGGCCCGCGAGCCGAGGCCATGGTGACAGAAAC	364
Db	741		CGCCGCGAGGGTGTGAGCGCGCGTGGCGGCCCGCGAGCCGAGGCCATGGTGACAGAAAC	800
Qy	365		CAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCGCCGG	424
Db	801		CAACAATGCCGAGAACACGGAAGCGCTGCTGGCCGGCGAGAGCTCGGACTCGGGCGCCGG	860
Qy	425		CCTCGAGCTGGGAATCGCCTCTCTCCCCACGCCCGGCTCCACCGCTTCCAGGGCGGCAA	484
Db	861		CCTCGAGCTGGGAATCGCCTCTCTCCCCACGCCCGGCTCCACCGCTTCCAGGGCGGCAA	920
Qy	485		GGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGAACGC	544
Db	921		GGCCGACGACCCGAGCTGGTGCAAGACCCGAGTGGGCACATCAAGCGACCCATGAACGC	980
Qy	545		CTTCATGGTGTGGTTCGAGATCGAGCGCGCAAGATCATGGAGCAGTCGCCGCATGCA	604
Db	981		CTTCATGGTGTGGTTCGAGATCGAGCGCGCAAGATCATGGAGCAGTCGCCGCATGCA	1040
Qy	605		CAACGCCGAGATCTCCAAGCGGCTGGGCAAAACGCTGGAAGCTGCTCAAAGACAGCGACAA	664
Db	1041		CAACGCCGAGATCTCCAAGCGGCTGGGCAAAACGCTGGAAGCTGCTCAAAGACAGCGACAA	1100
Qy	665		GATCCCTTTTCATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGCTGACTACCCCGA	724
Db	1101		GATCCCTTTTCATTTCGAGAGGCGGAGCGGCTGCGCCTCAAGCACATGGCTGACTACCCCGA	1160
Qy	725		CTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCTCTCGGC	784
Db	1161		CTACAAGTACCGGCCAGGAAGAAGGTGAAGTCCGGCAACGCCAACTCCAGCTCTCTCGGC	1220
Qy	785		CGCCGCCTCTCCAAGCGGGGGAGAGGGAGACAAGTTCGTTGGCAGTGGCGGGGGCGG	844
Db	1221		CGCCGCCTCTCTCCAAGCGGGGGAGAGGGAGACAAGTTCGTTGGCAGTGGCGGGGGCGG	1280
Qy	845		CCATGGGGGCGGCGGCGCGGCGGGAGCAGCAACCGGGGGGAGGAGGCGCGGTGCGAG	904
Db	1281		CCATGGGGGCGGCGGCGGCGGGAGCAGCAACCGGGGGGAGGAGGCGCGGTGCGAG	1340
Qy	905		TGGCGGCGGCGGCAACTCCAACCGGCGCAGAAAAAGAGCTCGGCTCCAAGTGGCGGG	964
Db	1341		TGGCGGCGGCGGCAACTCCAACCGGCGCAGAAAAAGAGCTCGGCTCCAAGTGGCGGG	1400
Qy	965		CGGCGCGGGCGGTGGGGTTAGCAAAACCGCACGCCAAGCTCATCTGGCAGCGCGCGCGG	1024
Db	1401		CGGCGCGGGCGGTGGGGTTAGCAAAACCGCACGCCAAGCTCATCTGGCAGCGCGCGCGG	1460

Qy	1025	CGGCGGGAAGACAGCGGCTGCCGCCGCCGCCCTCCTTCGCCGCCGAACAGCGGGGGCCGC	1084
Db	1461	CGGCGGGAAGACAGCGGCTGCCGCCGCCGCCCTCCTTCGCCGCCGAACAGCGGGGGCCGC	1520
Qy	1085	CGCCCTGCTGCCCTTGGGCGCCGCCGCCGACCACCTCGCTGTACAAAGCGCGGACTCC	1144
Db	1521	CGCCCTGCTGCCCTTGGGCGCCGCCGCCGACCACCTCGCTGTACAAAGCGCGGACTCC	1580
Qy	1145	CAGCGCTTCGGCTCCGCCCTCCTCGGCAGCCTCGGCCTCCGCAGCGCTCGCGGCCCCGGG	1204
Db	1581	CAGCGCTTCGGCTCCGCCCTCCTCGGCAGCCTCGGCCTCCGCAGCGCTCGCGGCCCCGGG	1640
Qy	1205	CAAGCACCTTGGCGGAGAAGAAGGTGAAGCGCGTCTACCTGTTTCGGCGGCCCTGGGCAGTCT	1264
Db	1641	CAAGCACCTTGGCGGAGAAGAAGGTGAAGCGCGTCTACCTGTTTCGGCGGCCCTGGGCAGTCT	1700
Qy	1265	GTCGTTCGCCCGTGGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGCTGTA	1324
Db	1701	GTCGTTCGCCCGTGGGCGGCGTGGGCGCGGGAGCCGACCCAGCGACCCCTGGGCTGTA	1760
Qy	1325	CGAGGAGGAGGGCGCGGGCTGCTCGCCCGACGCGCCAGCCTGAGCGGGCCGAGCAGCGC	1384
Db	1761	CGAGGAGGAGGGCGCGGGCTGCTCGCCCGACGCGCCAGCCTGAGCGGGCCGAGCAGCGC	1820
Qy	1385	CGCCTCGTCCCCCGCCGCGCGGCGCTCGCCCGCCGACCACCGCGGTACGCCAGCCTGCG	1444
Db	1821	CGCCTCGTCCCCCGCCGCGCGGCGCTCGCCCGCCGACCACCGCGGTACGCCAGCCTGCG	1880
Qy	1445	CGCGCCTTCGCCCGCCCGCTCCAGCGCGCCCTCGCAGCGTCTCTCTCGGCTCGTCCCA	1504
Db	1881	CGCGCCTTCGCCCGCCCGCTCCAGCGCGCCCTCGCAGCGTCTCTCTCGGCTCGTCCCA	1940
Qy	1505	CTCTCTCTCTCTCTCTCTCTCGGCTCTCTCGTCTCCGACGACGAGTTCGAAGACGACCT	1564
Db	1941	CTCTCTCTCTCTCTCTCTCTCGGCTCTCTCGTCTCCGACGACGAGTTCGAAGACGACCT	2000
Qy	1565	GCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCGAGTTCACTTCGTC	1624
Db	2001	GCTCGACCTGAACCCAGCTCAAACCTTTGAGAGCATGTCCCTGGGCGAGTTCACTTCGTC	2060
Qy	1625	GTCGGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACTTCGA	1684
Db	2061	GTCGGCGCTCGACCGGACCTGGATTTTAACTTCGAGCCCGGCTCCGGCTCGCACTTCGA	2120
Qy	1685	GTTCCCGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCGAGTC	1744
Db	2121	GTTCCCGGACTACTGCACGCCGAGGTGAGCGAGATGATCTCGGGAGACTGGCTCGAGTC	2180
Qy	1745	CAGCATCTCCAACCTGGTTTTCACCTACTGAAGGGCGCGCAGGCAGGGAGAAGGGCCGGG	1804
Db	2181	CAGCATCTCCAACCTGGTTTTCACCTACTGAAGGGCGCGCAGGCAGGGAGAAGGGCCGGG	2240
Qy	1805	GGGGGTAGGAGAGGAGAAAAAAAGTGAATAAAAGAAACGAAAAGGACAGACGAAGAGT	1864
Db	2241	GGGGGTAGGAGAGGAGAAAAAAAGTGAATAAAAGAAACGAAAAGGACAGACGAAGAGT	2300
Qy	1865	TTAAAGAGAAAAAGGAAAAAAGAAAAAAGTAAGCAGGGCTCGTTCCGCCCGGTTCTC	1924
Db	2301	TTAAAGAGAAAAAGGAAAAAAGAAAAAAGTAAGCAGGGCTCGTTCCGCCCGGTTCTC	2360
Qy	1925	GTCGTTCGATCAAGGAGCGCGGCGGCGTTTGGACCCGCGCTCCCATCCCCACCTTCCC	1984
Db	2361	GTCGTTCGATCAAGGAGCGCGGCGGCGTTTGGACCCGCGCTCCCATCCCCACCTTCCC	2420
Qy	1985	GGGCGGGGACCCACTCTGCCAGCCGAGGGAGCGCGAGGAGGAAGAGGGTAGACAGGG	2044
Db	2421	GGGCGGGGACCCACTCTGCCAGCCGAGGGAGCGCGAGGAGGAAGAGGGTAGACAGGG	2480

Qy	2045	GCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAAGCGACTTCG	2104
Db	2481	GCGACCTGTGATTGTTGTTATTGATGTTGTTGTTGATGGCAAAAAAAAAAAGCGACTTCG	2540
Qy	2105	AGTTTGCTCCCTTTGCTTGAAGAGACCCCCCTCCCTTCCAACGAGCTTCCGGAAGTGT	2164
Db	2541	AGTTTGCTCCCTTTGCTTGAAGAGACCCCCCTCCCTTCCAACGAGCTTCCGGAAGTGT	2600
Qy	2165	CTGCACCCCCAGCAAGAAGCGAGTTAGTTTCTAGAGACTTGAAGGAGTCTCCCTTC	2224
Db	2601	CTGCACCCCCAGCAAGAAGCGAGTTAGTTTCTAGAGACTTGAAGGAGTCTCCCTTC	2660
Qy	2225	CTGCATCACCACCTTGTTTGTGTTTATTGCTTCTTGCTCAAGAAAGGAGGGGAGAAC	2284
Db	2661	CTGCATCACCACCTTGTTTGTGTTTATTGCTTCTTGCTCAAGAAAGGAGGGGAGAAC	2720
Qy	2285	CCAGCGCACCCCTCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGGGGTG	2344
Db	2721	CCAGCGCACCCCTCCCCCTTTTTTAAACGCGTGATGAAGACAGAAGGCTCCGGGGTG	2780
Qy	2345	ACGAATTTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCACGCAG	2404
Db	2781	ACGAATTTGGCCGATGGCAGATGTTTGGGGGAACGCCGGGACTGAGAGACTCCACGCAG	2840
Qy	2405	GCGAATTCCTGTTGGGGCCTTTTTTCTCCCTCTTTTCCCTTGCCCCCTCTGCAGCC	2464
Db	2841	GCGAATTCCTGTTGGGGCCTTTTTTCTCCCTCTTTTCCCTTGCCCCCTCTGCAGCC	2900
Qy	2465	GGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTACCGGCGCTAGGAAATGACCC	2524
Db	2901	GGAGGAGGAGATGTTGAGGGGAGGAGGCCAGCCAGTGTACCGGCGCTAGGAAATGACCC	2960
Qy	2525	GAGAACCCCGTTGGAAGCGCAGCAGCGGAGCTAGGGGCGGGGCGGAGGAGACACGAA	2584
Db	2961	GAGAACCCCGTTGGAAGCGCAGCAGCGGAGCTAGGGGCGGGGCGGAGGAGACACGAA	3020
Qy	2585	CTGGAAGGGGTTTCAGGTCAAACTGAAATGGATTGCACGTTGGGGAGCTGGCGGCGGC	2644
Db	3021	CTGGAAGGGGTTTCAGGTCAAACTGAAATGGATTGCACGTTGGGGAGCTGGCGGCGGC	3080
Qy	2645	GGCTGCTGGGCTCCGCCCTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCAGACC	2704
Db	3081	GGCTGCTGGGCTCCGCCCTCTTTTCTACGTGAAATCAGTGAGGTGAGACTTCCAGACC	3140
Qy	2705	CCGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGAGTCACTGGAGGGCG	2764
Db	3141	CCGAGGCGTGGAGGAGAGGAGACTGTTTGATGTGGTACAGGGGAGTCACTGGAGGGCG	3200
Qy	2765	AGTGGTTTCGGAAAAAAAAAAGAAAAAAAAAGGG	2797
Db	3201	AGTGGTTTCGGAAAAAAAAAAGAAAAAAAAAGGG	3233

RESULT 13

ABV22264

ID ABV22264 standard; cDNA; 4467 BP.

XX

AC ABV22264;

XX

DT 13-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 22255.

XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW

pharmacogenomic marker; gene; ss.

XX